Spatiotemporal machine learning on scanner-free body surface potential imaging aided by multiscale modeling for personalized atrial fibrillation treatment

Yingjing Feng

université de BORDEAUX

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR
DE L'UNIVERSITÉ DE BORDEAUX

ECOLE DOCTORALE MATHÉMATIQUES ET INFORMATIQUE

CALCUL SCIENTIFIQUE ET MODÉLISATION

Par **Yingjing FENG**

Spatiotemporal machine learning on scanner-free body surface potential imaging aided by multiscale modeling for personalized atrial fibrillation treatment

Sous la direction de : **Edward VIGMOND**

Soutenue le 14 décembre 2021

Membres du jury :

M. Gernot PLANK	Professeur	Medical University of Graz	Rapporteur, Président
Mme. Molly MALECKAR	Professeure	Simula Research Laboratory	Rapporteure
M. Frédéric SACHER	Professeur (PU-PH)	Université de Bordeaux	Examinateur
Mme. Maria GUILLEM	Professeure Associée	Polytechnic University of Valencia	Examinatrice
M. Edward VIGMOND	Chargé de Recherche	Université de Bordeaux	Directeur

Apprentissage automatique spatiotemporel sur l'imagerie du potentiel de surface corporelle sans scanner aidé par la modélisation multi-échelle pour la personnalisation du traitement de la fibrillation auriculaire

Résumé: La fibrillation auriculaire (FA) est une activation rapide et irrégulière des oreillettes. Elle est l'arythmie clinique la plus courante. Le premier traitement préconisé est l'ablation par cathéter. Cette procédure est la plus efficace et améliore la qualité de vie. Néanmoins les protocoles standards utilisés démontrent un succès à long terme sous-optimal, justifiant la nécessité de mettre en place une ablation personnalisée. Les cartographies des potentiels à la surface corporelle (CPSC) sont des signaux électriques enregistrés sur le torse par un réseau d'électrodes, fournissant ainsi une vue panoramique non-invasive des oreillettes. Il est cependant difficile d'interpréter les CPSC pour le traitement de la FA. Cette thèse vise à développer de nouveaux algorithmes d'apprentissage automatique pour interpréter les CPSC afin de proposer un traitement personnalisé de la FA, sans faire appel à la résolution du problème inverse. À l'aide d'une modélisation multi-échelle un ensemble de données synthétiques a été généré. Il est composé d'environ 3000 épisodes de FA détaillés, ainsi que les CPSC à 252 dérivations et les électrocardiogrammes (ECG) à 12 dérivations, représentatifs des patients atteints de FA. Cet ensemble constitue notre base de données d'entraînement pour le développement des algorithmes et il permet d'expliquer les bases biophysiques de la prédiction des résultats pour l'ablation focale. Les algorithmes suivants ont été développés :

1) En combinant des sources périodiques auriculaires (SPA) équivalentes extraites de CPSC ou d'ECG avec un classificateur de forêt aléatoire, nous avons pu distinguer une source focale (SF) d'un rotor. Ceci nous permet de déduire si une SF peut initier une FA durable, ainsi que de déterminer la chambre auriculaire de la SF, avec une robuste de l'algorithme à la variabilité inter-patients sur l'ensemble de données synthétiques. L'analyse rétrospective des CPSC préopératoires a montré que les patients atteints de FA paroxystique, avec des épisodes de FA classés comme étant entraînés par une SF provenant d'une seule chambre auriculaire par notre algorithme, présentaient une récidive de FA inférieure jusqu'à 3 ans après ablation par cathéter au CHU de Bordeaux.

2) Le spectre des SPA, une fonction de la valeur maximale d'autocorrélation du SPA sur la durée du cycle, a été ensuite développé via des processus gaussiens pour représenter l'état de FA des patients. L'analyse des CPSC préopératoires de patients souffrant d'une FA persistante a révélé qu'une autocorrélation maximale plus élevée sur des durées de cycle de 220 à 230 ms était associée à une récidive de FA plus élevée jusqu'à 4 ans après ablation par cathéter au CHU de Bordeaux.

3) Un réseau de neurones récurrents convolutifs de graphe a également été développé afin de détecter le début de l'activation centrifuge (ATC) à partir des CPSC. Les ATC incluent non seulement des ATC actifs provenant d'une SF, mais également des ATC passifs provenant de la chambre auriculaire opposée par connexions inter-auriculaires (faisceau de Bachmann, foramen ovale et sinus coronaire). La détection par ATC a amélioré la détermination non-invasive des mécanismes de FA, incluant les SF et les autres sources de FA provenant d'une chambre unique, et, pour la première fois, les macro-réentrées circulant via les connexions inter-auriculaires.

Ces trois algorithmes permettent de déduire les mécanismes de la FA à partir des CPSC sans faire appel à l'imagerie cardiaque. Ces méthodes suggèrent des cibles d'ablation afin d'améliorer les résultats du traitement et élargissent le champ actuel des méthodes de cartographie non-invasive, permettant de faire progresser le diagnostic préopératoire non-invasif pour la personnalisation du traitement de la FA.

Mots-clés: fibrillation auriculaire, cartographie du potentiel de surface corporelle, ablation par cathéter, apprentissage profond géométrique, apprentissage automatique, jumeau numérique

Spatiotemporal machine learning on scanner-free body surface potential imaging aided by multiscale modeling for personalized atrial fibrillation treatment

Abstract: Atrial fibrillation (AF), the rapid and irregular activation of the atria, is the most common clinical arrhythmia. Catheter ablation therapy is the most effective treatment and improves the quality of life, but standard protocols show sub-optimal long-term success, substantiating the need for personalized ablation. Body surface potential maps (BSPMs) are electrical signals recorded on the torso by an array of electrodes, which non-invasively provide a panoramic view of atria, but it is challenging to interpret BSPMs for AF treatment. This thesis sought to develop novel machine learning algorithms to analyze BSPMs to inform personalized AF treatment, without solving the inverse problem. Using multiscale modeling, a synthetic dataset of nearly 3000 detailed AF episodes along with 252-lead BSPMs and 12-lead electrocardiograms (ECGs), representative of AF patients, was generated. This provided training data for algorithm development, and elucidated the biophysical bases of outcome prediction for focal ablation. The following algorithms were developed:

1) Combining equivalent atrial periodic sources (APSs) extracted from BSPMs or ECGs with a random forest classifier, we were able to discriminate between a focal source (FS) and a rotor, infer whether an FS can initiate sustained AF, and determine the chamber of FS, with robustness to inter-patient variability on the synthetic dataset. Retrospective analysis of pre-operative BSPMs showed paroxysmal AF patients with AF episodes classified as driven by an FS from a single atrial chamber by our algorithm had lower AF recurrence up to 3 years following catheter ablation from the CHU Bordeaux.

2) The APS spectrum, a function of the maximal autocorrelation value of APS over cycle length, was subsequently developed with Gaussian processes to represent the AF condition of patients. Analysis of pre-operative BSPMs of persistent AF patients revealed that a higher maximal autocorrelation over cycle lengths from 220 to 230 ms was associated with a higher AF recurrence up to 4 years following catheter ablation from the CHU Bordeaux.

3) A graph convolutional recurrent neural network was also developed to detect the onset of centrifugal activation (CAT) from BSPMs. The CATs include not only active CATs originating from an FS, but also passive CATs originating from the opposite atrial chamber through discrete inter-atrial connections (Bachmann's bundle, fossa ovalis, and the coronary sinus). CAT detection improved the non-invasive determination of AF mechanisms, including FS and other single-chamber drivers, and, for the first time, macro-reentries circulating through inter-atrial connections.

All three algorithms inferred AF mechanisms from BSPMs without cardiac imaging. The methods suggest ablation targets for better treatment outcomes, and broaden the scope of current non-invasive mapping methods, which together advance preoperative non-invasive diagnosis for personalized AF treatment.

Keywords: atrial fibrillation, body surface potential map, catheter ablation, geometric deep learning, machine learning, digital twin

---

Unité de recherche

Institut Hospitalo-Universitaire L'Institut de Rythmologie et modélisation Cardiaque (IHU-LIRYC),
F-33600 Pessac-Bordeaux, France

UMR 5251 - L'Institut de Mathématiques de Bordeaux (IMB),
Université de Bordeaux, 33000 Bordeaux, France

Substantial résumé en français

Apprentissage automatique spatiotemporel sur l'imagerie du potentiel de surface corporelle sans scanner aidé par la modélisation multi-échelle pour la personnalisation du traitement de la fibrillation auriculaire

La fibrillation atriale (FA) est l'arythmie soutenue la plus fréquente., Elle consiste en une activation rapide et irrégulière des oreillettes. L'ablation par radiofréquence constitue le traitement gold-standard de la FA. Elle se base sur l'isolement des veines pulmonaires (IVP) plus ou moins associées à l'ablation de substrat. Cependant, ces différents protocoles restent encore insuffisants pour prévenir les récurrences de la FA à long terme. C'est pourquoi il est nécessaire d'adapter un protocole d'ablation personnalisée pour chaque patient.

Les techniques de cartographie non invasives constituent un outil de planification préopératoire essentiel pour une ablation personnalisée. Les électrocardiogrammes (ECG), utilisés pour diagnostiquer les anomalies cardiaques depuis 1908, constituent l'outil de cartographie non invasif le plus important pour la FA. Les cartes de potentiel de surface corporelle (CSSP) enregistrées sur le torse par un maillage d'électrodes sont des signaux ECG de haute résolution qui fournissent en outre une vue panoramique des oreillettes de manière non invasive. L'imagerie ECG reconstruit l'activité auriculaire à l'aide de CSSP et de géométries du torse et du cœur spécifiques au patient en résolvant un problème inverse. Cependant, la reconstruction par imagerie ECG des modèles d'ondes d'activation complexes généralement observée chez les patients atteints de la FA n'a pas encore été validée. D'autres approches existent comme l'analyse spectrale et l'organisation des activités auriculaires, mais elles n'ont pas pu révéler les mécanismes spécifiques de la FA.

Par conséquent, nous avons cherché à développer des techniques avancées pour extraire les valeurs pronostiques sous-jacentes à partir des CSSP et des ECG des patients. Nous avons envisagé d'utiliser des techniques d'apprentissage automatique, car elles apprennent automatiquement la transformation des données en entrées en résultats de sortie. Cependant, la création d'une solution de bout en bout présente plusieurs défis. Le processus de collecte des données, comprenant à la fois la procédure d'ablation et la gestion des suivis post-ablation, est long et coûteux. Il en résulte une sensibilité à l'erreur importante dans la gestion des données cliniques. De plus, il n'existe actuellement pas de classification des activités auriculaires, car ce n'est pas une routine clinique d'enregistrer simultanément les signaux intracardiaques et les CSSP. Ensemble, ceux-ci ont limité le nombre de données de référence (telles que les résultats de l'ablation et les activités auriculaires) dans notre ensemble de données. De plus, le résultat du traitement d'un patient ne correspond plus à chaque signal collecté, il n'est donc plus approprié d'utiliser le signal et le résultat au niveau du patient comme paire d'entrées-sorties pour entraîner le modèle d'apprentissage automatique, connu sous le nom de

problème « classification multi-instance ».

Modélisation informatique du patient FA pour le développement d'outils de diagnostic

Pour résoudre ces problèmes d'application de l'apprentissage automatique aux CSSP et aux ECG, nous avons eu recours à une modélisation multi-échelle, qui fournit un ensemble de données synthétiques représentatif des patients atteints de FA. L'ensemble de données se compose de 2977 épisodes de FA enregistrées à l'aide de CSSP 252 électrodes et d'ECG 12 dérivations. Ces enregistrements obtenus à partir des épisodes de FA simulées ont permis d'obtenir des données de référence qui ont été utilisées pour la formation de modèles d'apprentissage automatique.

L'ensemble de données synthétiques a été construit à partir de cinq maillages auriculaires spécifiques aux patients, obtenus à partir d'imageries par résonance magnétique améliorées au gadolinium tardif, avec en moyenne 489 000 nœuds et 1,18 M d'éléments, qui ont fourni des géométries réalistes pour la simulation de l'activité auriculaire. Le comportement électrique de chaque nœud a été mis en œuvre par des modèles ioniques non linéaires avec des paramètres électriques spécifiques à la région, et la propagation des ondes a été modélisée dans une équation monodomaine, résolue à l'aide d'une méthode d'éléments finis sur le logiciel CARPentry. Les CSSP et les ECG ont été générés par modélisation prospective. Les épisodes de FA ont été simulés en introduisant des foyers ectopiques et/ou des rotors. Différent paramètres ont été modulées : la localisation des des foyers ectopiques et des « cores » des rotors, les fréquences des sources focales (SF), la taille des noyaux de rentrée et l'hétérogénéité électrique des oreillettes.

En plus de fournir des données d'entraînement pour les algorithmes d'apprentissage automatique, l'ensemble de données synthétiques a également élucidé les bases biophysiques de la prédiction de l'effet de l'ablation focale. Cela a été fait en simulant l'ablation localisée des épisodes induits par des SF. En fonction de l'inductibilité de la FA soutenue, nous avons classé la SF en deux types, « SF de type driver », où l'activation continue de la SF est essentielle pour conduire l'épisode de FA ; ainsi que les « SF de type initiateur », qui initie une FA auto-entretenue sans qu'il soit nécessaire d'être présent tout au long d'un épisode de FA. En comparant le comportement électrique entre ces deux types de SF, les rotors soutenus présents pendant les tirs d'ablation de SF sont plus susceptibles d'entraîner un épisode de FA soutenu après l'ablation localisée. La fréquence et la localisation de la source focale, ainsi que la présence de substrat, ont également été montrés comme des facteurs importants qui ont un impact sur la terminaison de la FA après la suppression des SF.

Sur le jeu de données synthétique et le jeu de données patient, nous avons ensuite

développé trois algorithmes. La mise en œuvre de chaque algorithme a été conçue pour s'adapter à son objectif d'apprentissage et à la disponibilité des données d'entraînement.

Algorithme 1. Détection non invasive de la source focale et du substrat arythmogène

Le premier algorithme a été conçu pour extraire la présence et la localisation de la SF, ainsi que pour prédire la terminaison de la FA après ablation localisé. . Comme les données CSSP FA de référence n'étaient disponibles que dans l'ensemble de données synthétiques, nous avons créé un modèle pour prédire le mécanisme à l'aide de l'ensemble de données synthétiques et appliqué le modèle à l'ensemble de données cliniques.

Nous avons d'abord extrait plusieurs sources périodiques auriculaires équivalentes (SPA) à partir de CSSP ou d'ECG, en utilisant la technique de séparation aveugle de source. Cet algorithme suppose que les signaux enregistrés ont été générés à partir d'un mélange de plusieurs sources périodiques. Nous avons ensuite extrait la durée de cycle et la périodicité de chaque SPA, ainsi que la contribution de la source périodique la mieux classée à tous les signaux de CSSP (s_1-to-lead contribution) de ces SPA équivalents, afin qu'ils agissent comme des caractéristiques des potentiels de surface du corps. Ensuite, nous avons construits des classificateurs de forêts aléatoires pour mapper ces caractéristiques aux cibles de prédiction souhaitables.

Le résultat a démontré une grande précision sur différents maillages de patients sur l'ensemble de données synthétiques. En particulier, la caractéristique de la contribution s_1-to-lead a démontré un modèle spatial cohérent pour les FS provenant de localisations similaires sur les oreillettes, à travers différents maillages de patients, et divers degrés d'hétérogénéité électrique dans les oreillettes. Cela a été fait sans fournir d'informations spatiales, telles que les localisations des électrodes CSSP. L'identification de la présence de SF et la prédiction de la durabilité de la FA étaient également robustes à une variation extrême de l'emplacement du gilet (jusqu'à 10 cm) et de la rotation (jusqu'à 20 degrés).

Sur l'ensemble de données des patients du CHU de Bordeaux, l'analyse rétrospective des CSSP préopératoires a démontré que les patients atteints de FA paroxystique avec des épisodes de FA classés comme entraînés par une SF d'une seule chambre auriculaire par notre algorithme présentaient une récidive de FA inférieure jusqu'à 3 ans après l'ablation par cathéter.

Algorithme 2. Spectre de source périodique auriculaire non invasif pour la représentation des conditions de la FA

Notre deuxième algorithme était destiné à prédire le résultat de l'ablation à long terme, directement à partir des ECG ou des CSSP du patient. Par conséquent, notre modèle a été conçu pour être simple et interprétable.

Un spectre SPA non invasif (SSPA), qui est fonction de la valeur d'autocorrélation maximale de l'SPA sur la durée du cycle, a ensuite été développé à l'aide d'un modèle de régression de processus gaussien. L'SSPA non invasif était un modèle de substitution pour représenter la condition de FA, et peut être agrégé entre différents signaux d'un patient. Pour identifier la caractéristique la plus prédictive de l'SSPA non invasive, le spectre a été sous-échantillonné pour produire des covariables pour un modèle de régression. Les covariables ont été utilisées pour prédire la probabilité de survie des suivis post-ablation à l'aide d'un modèle de risque proportionnel de Cox, qui a également sélectionné une limite de décision pour séparer deux groupes de patients avec probabilités distinctes de récidive à long terme de la FA.

La validité de l'SSPA non invasive a été testée en la comparant à l'SSPA invasive sur l'ensemble de données synthétiques. Pour cela, nous avons montré que l'SSPA non invasif produisait une couverture élevée et une petite erreur absolue par rapport à l'SSPA invasif, ainsi qu'une correspondance étroite sur la durée estimée du cycle dominant. Des modèles peuvent également être observés dans l'SSPA estimé à partir de différents signaux du même patient.

À l'aide d'une validation croisée d'un facteur 10, l'analyse rétrospective des CSSP préopératoires de patients atteints de la FA persistante a révélé qu'une autocorrélation maximale plus élevée sur des durées de cycle de 220 à 230 ms était associée à une récidive de FA plus élevée jusqu'à 4 ans après l'ablation par cathéter.

Algorithme 3. Détection non invasive de l'activation centrifuge active et passive avec apprentissage en profondeur

Le troisième projet a utilisé l'apprentissage profond « deep learning » géométrique pour détecter le début de l'activation centrifuge (ATC) à partir des CSSP, ce qui a aidé à l'extraction de plusieurs mécanismes de FA qui étaient auparavant inaccessibles par l'imagerie ECG. Les ATC incluent non seulement les ATC actifs provenant d'une SF, mais également les ATC passifs provenant de la chambre auriculaire opposée par des connexions inter-auriculaires discrètes (faisceau de Bachmann, fosse ovale et sinus coronaire).

Lorsque nous avons estimé la probabilité de l'ATC pour chaque milliseconde, les données d'apprentissage ont été augmentées de la longueur (en millisecondes) de la

série chronologique. Cela nous a permis d'utiliser des modèles d'apprentissage automatique plus flexibles tels que des modèles de deep learning. Un réseau de neurones récurrents à convolution graphique a été développé pour détecter le début de l'ATC à partir des CSSP. Le modèle comprend : (1) une couche EdgeConv en tant qu'encodeur spatio-temporel de CSSP, qui contient un modèle de perceptron multicouche simple qui apprend une représentation temporelle condensée du signal de chaque nœud (c'est-à-dire électrode) à travers un message passant entre les voisins ; (2) un réseau neuronal récurrent bidirectionnel en tant qu'encodeur temporel, qui se compose de cellules de mémoire à long terme ; et enfin, (3) une couche de suréchantillonnage et une couche sigmoïde, qui ont suréchantillonné la sortie de l'encodeur temporel à sa résolution d'origine, et ont produit quatre séries temporelles probabilistes en tant que sorties, représentant la probabilité de l'ATC actifs et passifs sur deux chambres auriculaires à chaque temps.

Notre modèle est le premier algorithme à détecter le timing de l'événement à partir des CSSP. Comparé à un réseau de neurones récurrent convolutif traditionnel, le module EdgeConv a amélioré la précision de l'estimation de l'ATC ainsi que le mécanisme de la FA. La détection des ATC actifs a donné une grande précision et une aire sous la courbe de ROC élevée à une fenêtre temporelle étroite de 20 ms, ce qui a permis d'améliorer la localisation SF par le premier algorithme. La détection des ATC passifs a donné une précision supérieure à 80% et près de 90% sous la courbe de ROC à une fenêtre de temps de 250 ms. La détection ATC a permis la détermination non invasive de plusieurs mécanismes de FA, y compris le FS et d'autres « drivers » à chambre unique, et, pour la première fois, des macro-réentrées circulant via des connexions inter-auriculaires.

Conclusion

Cette thèse visait à développer de nouveaux algorithmes d'apprentissage automatique pour analyser les CSSP, afin d'informer le traitement personnalisé de la FA.

La contribution la plus importante que nous ayons apportée a été le développement d'algorithmes d'apprentissage automatique spatio-temporel sur les CSSP, sans qu'il soit nécessaire de fournir des informations de localisation spécifiques comme données en entrée. Nous avons montré que l'extraction de la structure spatio-temporelle peut être effectuée en exploitant les statistiques communes de second ordre parmi les signaux de différents canaux, ou en intégrant la structure commune des électrodes voisines. Cela simplifie grandement les procédures cliniques, car les patients ne sont plus soumis au processus d'imagerie cardiaque. De plus, cela évite d'avoir à résoudre le problème inverse mal posé de l'imagerie ECG.

Les résultats des méthodes suggèrent des cibles d'ablation pour les patients, ce qui entraînera probablement de meilleurs résultats de traitement. Les mécanismes suggérés

par nos méthodes élargissent le champ des méthodes de cartographie non invasives actuelles. Ensemble, ces deux points forts de notre modèle font progresser le diagnostic préopératoire non invasif qui conduit à un traitement personnalisé de la FA plus accessible.

Acknowledgements

First of all, I would like to thank my PhD supervisor, Dr. Edward Vigmond, who led me to become an independent researcher, and especially, provided great suggestions in scientific writing. To my clinical supervisor, Dr. Mélèze Hocini, for providing invaluable clinical insights and resources that served as the bases of my projects. To my external supervisor, Prof. Steven Niederer, who provided invaluable insights to the science and life. To my collaborators of the atrial modeling project, Dr. Caroline Roney and Dr. Jason Bayer, for their guidance in atrial fibrillation modeling. To my PhD monitoring committee, Prof. Blanca Rodriguez, Prof. Esther Pueyo, Prof. Rémi Dubois and Dr. Mostafa Bendahmane, for their advice in my PhD progress and career. To my collaborators of patient outcome prediction projects, Carole Dumas-Pomier, Dr. Rozenn Mingam and Xavier Pillois for their kind help.

I would like to express my sincerest gratitude to my admirable colleagues, Dr. Jairo Rodriguez, Dr. Peter Langfield, Dr. Angel Moreno, Dr. Julien Bouyssier, and Dr. Mirabeau Sara, for being my role models and my dearest friends. I was also delighted to meet Dr. Marianna Meo, Dr. Hassan Adam Mahamat, Dr. Nolwenn Tan, Dr. Girish Ramlugun, Elodie Surget, Dr. Vladimír Sobota, Dr. Marta Nunez Garcia, and Sylvain Caubet, with whom I have spent wonderful time in the same institute. To the staff of the functional team in IHU-LIRYC, especially Sophie Panteix, Delphine Ge, Christiane Andriamandroso, Loïc Durand, and Ruben Yoyotte, from my research unit, IHU-LIRYC, I appreciate their kindness and supports for the execution of my PhD study.

My PhD study has been funded from the European Union Horizon 2020 Research and Innovation programme “Personalised In-silico Cardiology (PIC)” under the Marie Skłodowska-Curie grant agreement no. 764738. For which, I am very grateful, and especially, I would like to thank the coordinator of consortium, Prof. Pablo Lamata, for his humour and dedication to our career growth. It was also my pleasure to meet the other 15 fellows under the same project, Hongxing Luo, Maciej Marciniak, Cristobal Rodero, Jorge Corral-Acero, Francesca Margara, Filip Loncaric, Andrew Gilbert, Joao F Fernandes, Hassaan A Bukhari, Ali Wajdan, Manuel Villegas Martinez, Mariana Sousa Santos, Mehrdad Shamohammdi, Philip Westphal, and Valeria Galli in the consortium, for their support and company during my PhD journey. Some of them also have achieved their PhD qualification to-date.

Last but not the least, I would like to dedicate this thesis to my parents, and my best friends in Bordeaux, including Petya Stoeva, Yiye Jiang, Dr. Liangying Zhu, Dr. Chenhao Feng, Wenjing Ma, and Yanfei Xiang, and my best friend and partner, Robert Daniels.

“In solis sis tibi turba locis.” by Tibullus

Contents

Abbreviations

3D Three-Dimensional

ACF Auto-Correlation Function

ACh Acetylcholine

AF Atrial Fibrillation

AP Action Potential

APD Action Potential Duration

APSS Atrial Periodic Source Spectrum

AUROC Area Under the Receiver Operating Characteristic Curve

BSS Blind Source Separation

CAT Centrifugal Activation

CFAE Complex Fractionated Atrial Electrograms

CL Cycle Length

CMR Cardiac Magnetic Resonance

CNN Convolutional Neural Network

CRN Courtemache-Ramirez-Nattel

CRNN Convolutional Recurrent Neural Network

CT Computed Tomography

DF Dominant Frequency

DFS Driver-type Focal Source

ECG Electrocardiogram

ECGi Electrocardiogram Imaging

ERP Effective Refractory Period

FAF Atrial Fibrillation Induced by Focal Source

FFT Fast Fourier Transform

FS Focal Source

GCRNN Graph Convolutional Recurrent Neural Network

GNN Graph Neural Network

GP Gaussian Process

GPR Gaussian Process for Regression

ICA Independent Component Analysis

IFS Initiator-type Focal Source

LA Left Atrium

LGE-MRI Late Gadolinium-Enhanced Magnetic Resonance Imaging

LSTM Long Short-Term Memory

MAE Mean Absolute Error

MaxAC Maximum of Auto-Correlation Function

MLP Multi-Layer Perceptron

MRI Magnetic Resonance Imaging

NDI Nondipolar Component Index

PCA Principal Components Analysis

PS Phase Singularity Point

PV Pulmonary Vein

RA Right Atrium

RAF Reentrant-type Atrial Fibrillation

RNN Recurrent Neural Network

SO-BSS Second-Order Blind Source Separation

BSPM Body Surface Potential Map or Body Surface Potential Mapping

UAC Universal Atrial Coordinate

VCG Vectorcardiography

List of Figures

List of Tables

Chapter 1

Introduction

Atrial fibrillation (AF), a heart condition in which the atria activate rapidly and irregularly, is the most common arrhythmia seen clinically, affecting up to 9% of the population over the age of 65 [1]. The prevalence of AF is growing, likely due to increased longevity, as well as increased risks in the general population for related cardiovascular diseases, such as stroke, hypertension and obesity, as well as higher accessibility to screening [2, 1].

AF is a multifaceted and progressive disease [2], with variability between patients, as well as within the same patient over different disease stages. Catheter ablation therapy, which isolates or destroys parts of the cardiac tissue, is the most effective treatment for AF, improving quality of life and maintaining sinus rhythm longer than antiarrhythmic drugs [3]. It is however challenging to select the optimal ablation strategy for a patient.

Pulmonary vein (PV) isolation [4], which electrically isolates the PVs, is the most adopted ablation strategy, especially for the first procedure [5]. However, AF recurrence rates for PV isolation are about 40–50% within one year worldwide; typical long-term AF-free rates (after 3 to 5 years) are 54.1% paroxysmal AF and 41.8% non-paroxysmal AF according to a meta-analysis study [6]. In another multicenter study involving 589 patients, AF recurrence rates range between 40% to 60% one-year post-ablation, and further substrate or linear ablation did not significantly improve results [7]. Although reconnection across ablation lines in the PVs is likely a cause for late AF recurrence within the first year after the procedure [5], the association between the lower risk of AF recurrence post-ablation and durable PV isolation was found to be only modest [8]. There was also no significant difference between different standardized ablation approaches, i.e. PV isolation, PV isolation with linear ablation and PV isolation with complex fractionated atrial electrogram ablation [7]. These findings suggest that the AF treatment can potentially be improved by personalized ablation strategies.

Accurate non-invasive mapping technologies of AF patients for mechanism-directed ablation are still lacking [9]. Non-invasive electrophysiological signals, such as electrocardiograms (ECGs), have been used since 1908 [10] to diagnose cardiac abnormalities, including the diagnosis of AF in which f-waves are present as rapid wavelets with irregular length. Body surface potential maps (BSPMs), which acquire electrical potentials from the whole torso surface using a large array of electrodes, is thus far, the only technique providing a panoramic view of both atria, a significant advantage over intracardiac signals acquired sequentially, using a mapping catheter. However, these recordings are the surface projection of mixtures of intracardiac signals, where the potentials from the intracardiac sources diminish as the distance to epicardium increases, which makes the non-invasive interpretation of the complex activation patterns during AF episodes especially challenging.

The state-of-the-art non-invasive mapping techniques can be categorized into ECG imaging (ECGi), spectral analysis and organization measures. ECGi reconstructs atrial activity using BSPM signals, and patient-specific geometries of the torso and the heart, by solving an inverse problem. However, reconstruction of complex wave patterns by ECGi typically observed in AF patients, such as multiple breakthroughs, has not yet been validated [11]. A representative work of spectral analysis comes from Guillem et al. [12], which showed that the sites with dominant frequency on the BSPM are spatially correlated with dominant frequency sites mapped invasively. However, dominant frequency analyses are agnostic to the type of AF drivers. Measures on the organization of atrial activities have also been frequently used, including organization index [13], f-wave amplitude [14], and non-dipolar index [15], but none of these measures reveals the underlying AF mechanisms.

Machine learning algorithms automatically approximate an unknown mapping function from the input to the target output by being presented with many examples [16]. In recent years, deep learning algorithms have also successfully extracted spatiotemporal features in a wide range of applications, such as computer vision, language processing, and medical image analysis [17]. We sought to unveil the relationship between the treatment outcomes and the patient pre-operative signals via machine learning techniques. However, there exist several major challenges. The process to collect AF treatment outcomes is time-consuming and expensive, resulting in a limited number of AF patients enrolled. The treatment outcomes also do not necessarily correspond to each signal collected. For example, for a signal collected from an AF patient during sinus rhythm, it is inappropriate to categorize it as an AF episode. The many-to-one mapping between the input and the output is known as a problem of “multi-instance classification” [18], which complicates the training of machine learning models. Furthermore, we do not know the ground-truth atrial activity corresponding to the obtained BSPMs, as it is not a standard clinical routine to simultaneously record the intracardiac signals and BSPMs.

To solve these challenges, we resorted to multiscale atrial models of AF patients. They are computer models built on detailed atrial geometries made from patient meshes, with realistic electrophysiological parameters and computational rules, which are made to mimic the electrical behaviour of patient atria during AF episodes. When used properly, the simulations can act as "digital twins" of the AF patients, which not only yield mechanistic insights into the pathology for deductive reasoning, but also provide synthetic data and augment patient data to aid the development of machine learning models for inductive reasoning [19].

With computer models, a virtual cohort of an AF sub-type, with detailed views of both atrial activities and the body surface potentials, could be obtained at a much lower cost. On the other hand, the prediction targets of patient treatment outcomes, are only available from clinical followups. Therefore, the aim of this thesis is to develop novel machine learning techniques on BSPMs with interpretability and robustness, by exploiting the advantages of both synthetic data from multiscale modeling and patient data from clinical studies, in order to improve BSPM for personalized AF treatment. The prediction targets indicate AF mechanisms which can guide AF catheter ablation. As a proof-of-concept, the usefulness is demonstrated in prediction of long-term ablation outcomes from pre-operative patient signals. A brief overview of the organization of the chapters is included below.

- Chapter 2 - The background of the work, including the cardiac electrophysiology, AF mechanisms and treatment, multiscale atrial models, machine learning techniques, and related works.
- Chapter 3 - Methods and results for the generation of the synthetic dataset via multiscale modeling for Chapter 4, 5 and 6, with important insights that inspired the development of a machine learning algorithm presented in Chapter 4.
- Chapter 4 - Development of a machine learning algorithm for predicting AF mechanisms to separate reentrant from focal sources, and different types of focal sources, with an application of predicting late AF recurrence of paroxysmal AF patients post-ablation.
- Chapter 5 - Development of another algorithm to extract a patient cycle length spectrum metric, which acts as a surrogate model for the atrial state for AF patients, with application to predicting late AF recurrence of persistent AF patients post-ablation.
- Chapter 6 - Development of a novel deep-learning method to annotate active and passive centrifugal activations, which can indicate AF mechanisms including FS, single-atrial driver, and interatrial macroreentries that are inaccessible to ECGi.
- Chapter 7 - Discussion on the findings, impact, and future work.

Chapter 2

Background

Since the focus of our studies is on novel non-invasive mapping techniques for AF catheter ablation treatment, we first overview cardiac electrophysiology with a quantitative view. We then introduce the pathoelectrophysiology of AF mechanisms, the non-invasive mapping tools, ECGs and BSPMs, as well as the ablation therapy. In order to quantitatively analyze the non-invasive mapping signals, we introduce related machine learning and deep learning techniques, and then progress to the multiscale modeling which could aid such analysis. Finally, we present the related works for analyzing AF BSPMs, where we identified the need for developing robust machine learning methods for analyzing BSPMs to personalize AF treatment.

2.1 Basics of cardiac electrophysiology

The heart is an organ with a group of highly specialized muscle cells that contract and relax spontaneously throughout our life. Unlike other muscle cells, their excitation is only modulated by neurological stimulation. It is therefore vital to maintain a regular rhythm of contraction and relaxation, which ensures the transmission of oxygen, nutrients, metabolin and other substances to different organs of the body.

2.1.1 Action potential (AP)

The excitation of the cell is via the generation of an action potential (AP). The AP is a transient pulse of electrical charges across the cell membrane, where the transmembrane voltage V_m abruptly changes. The initiation of APs in the heart relies on a group of specialized cardiac cells, known as the "pacemaker cells", of the sinus-atrial node located on the top of the right atrium. The AP, once initiated, propagates to adjacent

cells. The process for AP propagating from sinoatrial node to the rest of the heart is shown in Fig. 2.1.

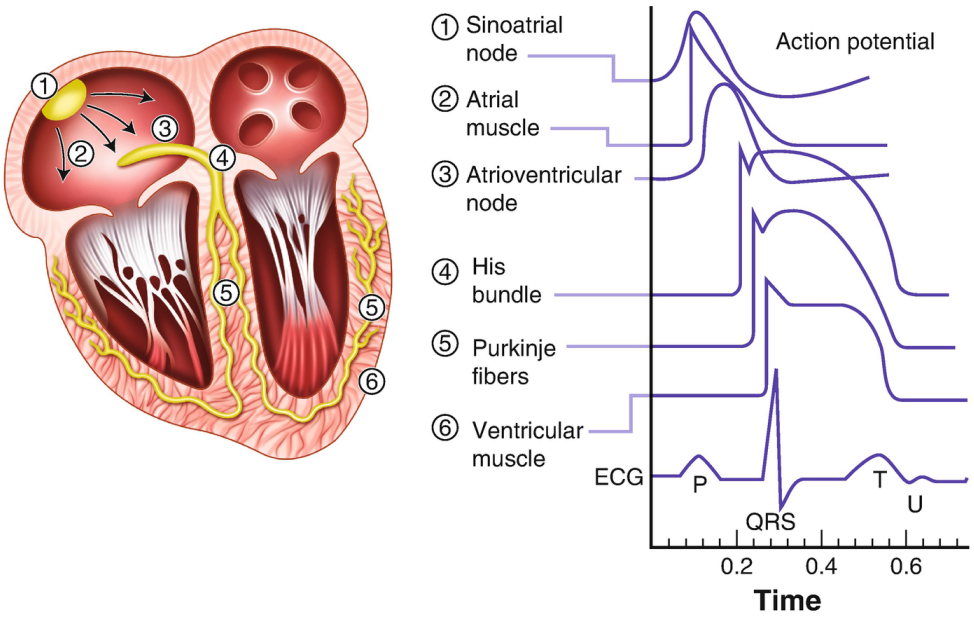


Fig. 2.1. The geometry of the heart, and the temporal relationship of activation as well as the morphologies of their APs in different parts of the heart. Figure reused with permission from Kusumoto et al. [20]

An AP is generated across the membrane of an excitable cell. The lipid bilayer of the membrane separates the interior from the outside of the cell. The movement of ions across the membrane via ion channels constitutes an electrical current.

The ions move from high to low concentration, which is known as diffusion. This process continues until the electrochemical equilibrium for that kind of ion is reached. The membrane electrical potential at that equilibrium is called "equilibrium potential" or "Nernst potential" (E_{ion}) [21]. Any deviation of the membrane voltage from the E_{ion} results in ion movement that restores towards E_{ion}, and in this sense, the membrane acts like a battery. The transmembrane voltage at the resting state, E_m, is the total sum of E_{ion} of all ions and is relatively constant, about -70 mV.

The thin lipid bilayer has a capacitance C_m [22]. On a patch of a membrane (Fig. 2.2), the transmembrane voltage, V_m, the ionic current is I_{ion}, the net membrane current is I_m, and they are related by

$$I_m = I_{ion} + C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} \tag{2.1}$$

To explain the abrupt change of transmembrane current during an AP, two physiologists, Hodgkin and Huxley, hypothesized in 1950s that the ionic currents are independent, and Na^+ and K^+ channels, while remaining closed normally, briefly open

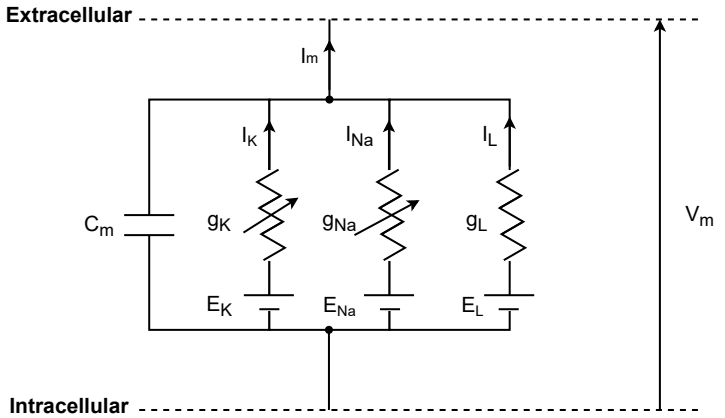


Fig. 2.2. A transmembrane patch modeled as parallel capacitive current and ionic currents in the Hodgkin-Huxley model. V_m: transmembrane voltage. g_K, g_{Na} and g_L: ionic conductance, with g_K and g_{Na} varying at transthreshold stage and static at resting stage of the AP. I_K, I_{Na} and I_L: potassium and sodium ionic currents and leak current. C_m: membrane capacitance.

during depolarization, which produces a greater depolarization [23]. They described the transmembrane current as Fig. 2.2, and provided the following equation

$$I_{ion} = g_K(V_m - E_K) + g_{Na}(V_m - E_{Na}) + g_L(V_m - E_L) \tag{2.2}$$

where g_K, g_{Na} and g_L are ionic conductance that change over time in a nonlinear manner. It can also be seen that the ionic currents depend on the voltage, which reflects the voltage-gating properties of the ion channels.

2.1.2 AP of atrial cells

The Hodgkin-Huxley model serves as an important foundation for quantitative electrophysiology, as we shall see next in cardiac AP models. The AP of a cardiac myocyte cell consists of rapid depolarization (phase 0), early repolarization (phase 1) plateau (phase 2), repolarization (phase 3) and resting (phase 4) phases. The electrical stimulus must exceed a threshold in order to initiate the AP. A stimulus with amplitude as a double of this threshold stimulus amplitude would not make a significant difference on the AP, although the latency between the stimulus and the depolarization is reduced [22]. When an AP is excited by a stimulus, the fast sodium channel opens, allowing a large amount of Na^+ influx, and triggers a sequence of concerted activities where ions, mainly Ca^{2+}, Na^+ and K^+, flow through the channels (Fig. 2.3).

The acetylcholine-activated potassium current $I_{K(ACh)}$ is activated by acetylcholine (ACh) released from the endings of vagal nerve. ACh release shortens action potential duration in the atrial myocytes, and increases the spatial heterogeneity of refractoriness in the atria, which further facilitates the reentry [25].

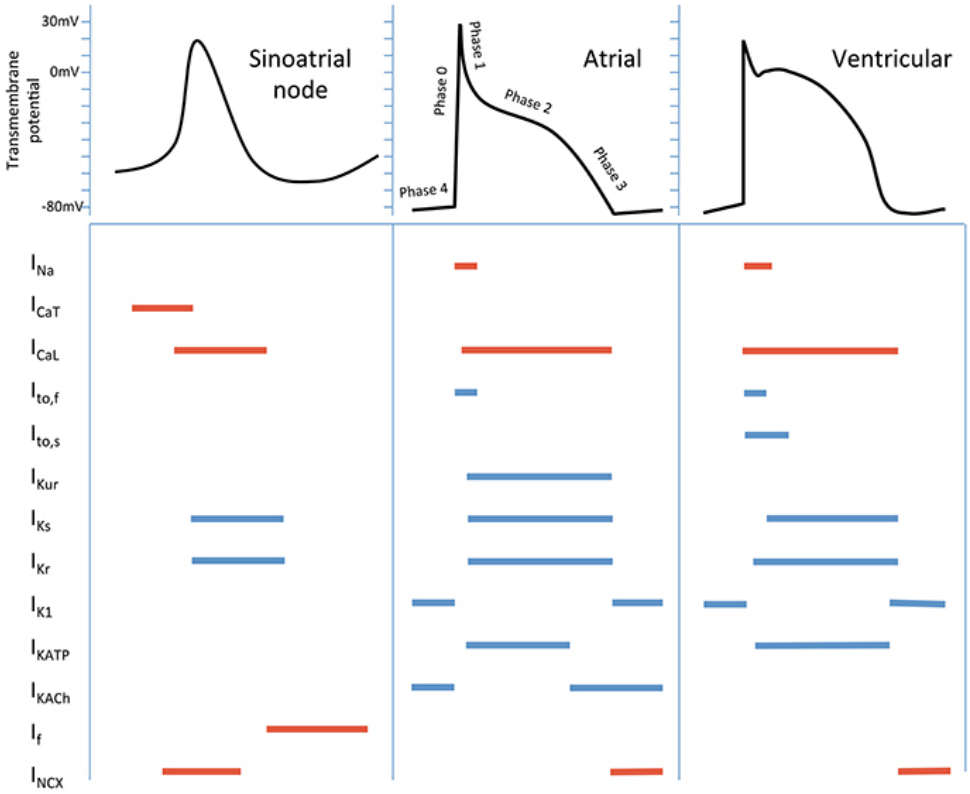


Fig. 2.3. Schematic representation of AP of a sinoatrial node, an atrial myocyte cell and a ventricular myocyte cell, with red showing inward currents and blue marking outward currents. I_{Na}: Na^+ current. I_{CaL} and I_{CaT}: L-type and T-type Ca^{2+} currents. $I_{to,f}$ and $I_{to,s}$: fast and slow transient outward K^+ currents. I_{Ks}, I_{Kr}, and I_{Kur}: slow, rapid and ultra-rapid components of the delayed rectifier K^+ current. I_{K1}: inward rectifier K^+ current. I_{KATP}: adenosine diphosphate-activated K^+ current. I_{KACh}: acetylcholine-activated K^+ current. I_f: "funny" current. I_{NCX}: Na^+/Ca^{2+} exchange current. Figure reused from Len et al. [24] under CC-BY license.

Several ionic models to describe the ionic current, I_{ion}, for atrial cells have been developed for the human atria, such as those by Nygren et al. [26], by Courtemache, Ramirez and Nattel (CRN) [27], by Maleckar et al. [28], by Grandi et al. [29] and by Koivumäki et al. [30]. These models emphasize different aspects of atrial arrhythmia, and so should be chosen according to specific applications [31]. Among these models, the CRN model was the only model to consistently reproduce beat-to-beat alternan [31], which is an important cellular electrical basis for the initiation of arrhythmia. The CRN model is also the most commonly used in tissue-scale AF modeling in the recent years [32].

2.1.3 Excitability of cell

During an AP, there is a period where the cell is non-excitable, which allows the relaxation of the cardiac myocytes, and serves as a protective mechanism against premature

atrial contraction and against multiple sequential stimuli. The effective refractory period (ERP), refers to the period when the cell cannot respond to a stimulus to depolarise. The AP duration (APD), the duration of the AP above a certain threshold, is often used as a surrogate for refractoriness. The value of this threshold is usually specified on a case-by-case basis.

2.1.4 Conduction velocity of depolarizing wave

Intracellular electrical excitation depends on the speed of the ions moving across the membrane, which is related to the driving force of the ions from the difference between E_{ion} and V_m, as shown in Eq. (2.2). Intercellular conduction is via the gap junctional channels. Gap junctions are low-resistance pathways, connecting the cytoplasm of two cells, both electrically and mechanically [33]. The intercellular conduction is anisotropic, where the conduction is faster along the longitudinal direction of the fiber, but slower in the transverse direction. High heterogeneity in the atria was also observed for the AP as well as the fiber orientation cross atrial regions [34, 35].

2.2 Atrial fibrillation (AF)

AF is a heart condition in which the atria activate very rapidly and irregularly. The exact mechanisms of AF are still not completely understood, and multiple hypotheses have been proposed. AF is also known to be a progressive disease, with both electrical and structural changes, rendering the atrial substrate more susceptible to longer AF episodes. Therefore, a variety of AF mechanisms have been found in patients, resulting in multiple ablation strategies targeting different mechanisms. The variety in the AF mechanisms has also posed challenges for the development of non-invasive mapping technologies.

2.2.1 AF mechanisms

Generally speaking, an AF episode requires an initiation mechanism and a perpetuation mechanism. An illustration of the basic AF mechanisms is summarized in Fig. 2.4, and the explanation is found in the main text followed.

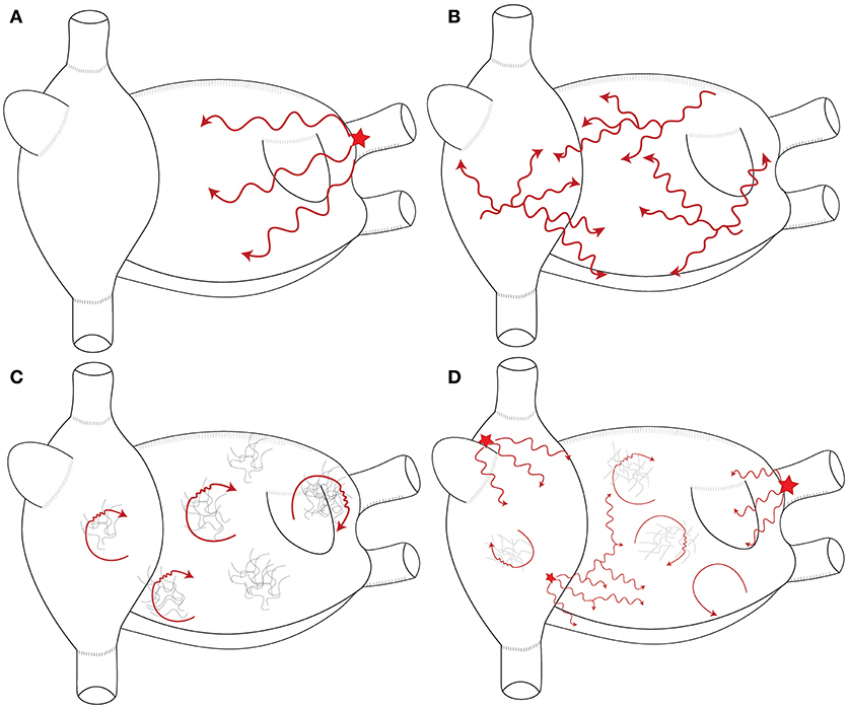


Fig. 2.4. An illustration of the basic AF mechanisms. A: Single focal trigger. B: Multiple wavelets that form and terminate, giving rise to daughter wavelets. C: Multiple reentries. D: Combination of focal triggers and reentries. Figure reused under license from Cheniti et al. [36].

AF initiation

An atrial ectopic beat is an electrical activation arising from a site other than the sinoatrial node. A ***focal source***, referring to a repetitively firing ectopic beat, can itself be regular, but can initiate fibrillatory activity with varied rates [37]. Reentry is initiated by an ectopic beat when the depolarizing wavefront encounters heterogeneously recovered tissue, causing the formation of a unidirectional conduction block [38]. The PV/LA junction with an APD gradient can act as a significant substrate for AF initiation. Spatial APD heterogeneity may also be attributed to the varied extent of parasympathetic activity across the atria, which triggers the ACh-activated potassium current $I_{K(ACh)}$ that shortens APD [25].

Triggers locations Ectopic beats are distributed primarily on the PVs, but are also common in non-PV regions. Haissaguerre et al. first reported the role of ectopic activity originating from PVs, which was found in 94% of the 45 AF patients with frequent AF episodes [4] (Fig. 2.4 A and D). The reported prevalence of non-PV ectopic beats in AF patients has been up to 60% [39]. Locations of non-PV foci include the LA and RA myocardium, LA appendage, inter-atrial septa, the vein of Marchall, the superior vena cava, and the coronary sinus [40, 41, 42, 43].

AF perpetuation

Multiple wavelets hypothesis The multiple wavelet hypothesis [44], developed by Moe in 1964, has been acknowledged as the main perpetuation mechanism for decades. In this hypothesis, the AF is believed to be sustained by multiple randomly meandering wavelets, which constantly break up and give birth to new independent wavelets [44] (Fig. 2.4 B). The clinical evidence for this hypothesis is the success of Cox-Maze procedure [45], which explains fibrillatory wavelets spontaneously terminating when they reach the unexcitable ablation lesions.

Localized source hypothesis However, the multiple wavelet hypothesis conflicts with the observation of dominant frequency (DF) sites clustering near sites such as the posterior LA near a PV ostium [46]. Using high-resolution optical mapping on sheep atria, it was also found that the phase singularity (PS) points were transient with a nonrandom spatial distribution, which likely resulted from the breakups of high-frequency organized waves [47]. These results together suggest that the wavelets are passive, and AF could be perpetuated from localized drivers. The "driver" term refers to the role of these localized mechanisms, which is akin to an "engine" driving the continuation of the AF episode. Multiple wavelets can appear at areas far away from these drivers, where wavebreaks lead to multiple wavefronts, so that identification of the underlying driver type is obscured.

The localized drivers consist of focal and/or reentrant activity (Fig. 2.4 D). The focal driver refers to a persistence focal source that perpetuates the AF episode, while ***reentry***, refers to the cardiac tissue being excited by circulation of the activation wavefront, around an anatomical or functional conduction obstacle [48] (Fig. 2.4 C).

Anatomical reentries rotate around scars or physical obstacles (such as the ring of triscupid valve in atrial flutter [49], or via the coronary sinus ostium AF [50]). For these reentries, the wavelength, the product of ERP and the average conduction velocity, should not exceed the circumference of the circuit in order to allow an excitable gap for the conduction front to advance. Micro-anatomical reentries (or intra-mural micro-reentries), as a kind of anatomical reentry, anchoring on intramural atrial bundles or highly fibrotic tissue, were observed using simultaneous optical mapping of endocardial and epicardial wall [51].

Functional reentry refers to ***spiral wave reentries***, or ***rotors***, where the core of a rotor remains unexcited, and the curvature of the depolarization wavefront modulates the conduction velocity [36]. Rotors have been observed in patients using basket catheters [52, 53].

Continuous spectrum The debate of these two hypotheses is still ongoing. For patients with self-terminating AF within 7 days, classified as "paroxysmal" AF patients, their AF episodes are believed to be mainly caused by trigger mechanisms. For both "persistent" AF patients, periods of AF lasting more than 7 days which can be terminated with cardioversion, and "permanent" AF patients, periods of AF longer than 7 days which cannot be terminated by treatment. As the disease progresses, the atria are found to go through a remodeling process which makes them more susceptible for AF perpetuation. Nonetheless, overlaps between the trigger and perpetuation mechanisms exist [38]. For example, persistent AF patients with minimal structural heart disease had one-year outcomes after PV isolation comparable to paroxysmal AF patients [54]. The current consensus is that the AF condition is described by a continuous spectrum, which, as the AF progresses, shifts from localized drivers to more diffuse complex wavelets perpetuated by an arrhythmogenic substrate [9].

2.2.2 Atrial remodeling

The progression of AF is accompanied by the process of atrial remodeling. The effects of atrial remodeling promoting AF perpetuation include shortening of APD, abnormal restitution, conduction slowing, and reduced automaticity or prolonged conduction of sino-atrial node [37, 55]. Contractile and structural remodeling include dilation in both atria [56], impaired atrial contractility, myocyte hypertrophy, myocyte death, and tissue fibrosis [38].

2.2.3 Standard ablation strategies of AF

Elimination of triggers

Since PV triggers are known to be the most common initiation mechanisms in AF patients, isolation of focal triggers by durable lesions is standardized clinical practice [5]. The isolation of triggers in the PVs, developed by Haïssaguerre et al. [4], is the standard procedure for the initial procedure of AF patients [5]. In order to reduce PV stenosis and prevent newly developed triggers from other PVs, an improved technique was developed which adds large circumferential ablation lesions on the LA to enclose all four PVs. This reduced operation time since there is no need for mapping before PV isolation. Non-PV ectopic beats were also found to be responsible for AF recurrence post-ablation, in both paroxysmal and persistent AF ablation [57], substantiating the need to map and ablate these triggers during procedures. Therefore, ablation of non-PV foci has also been taken into clinical practice [43].

Modification of substrate

Linear ablation The Cox-Maze procedure [45], creates a set of transmural lesions separating both atria into small compartments, with demonstrated success of AF-free maintenance for persistent AF patients. Variations of the procedure used a smaller lesion set, which led to linear atrial ablation strategies [38]. The linear ablation was usually performed in combination with other strategies such as PV isolation or circumferential LA ablation. The typical lines of ablation include the LA roof, LA posterior wall, the lateral mitral isthmus, and cavotricuspid isthmus [38]. However, as the lesions can lead to the development of LA flutter, the linear ablation strategy is not recommended for the initial ablation of paroxysmal AF patients [5].

Rotor core ablation Reentrant and focal drivers have been mapped and ablated with a 64-lead basket catheter. The initial studies [58] showed rotors are present in a large proportion of patients, but subsequent studies did not reproduce the same efficacy as the initial studies. ECGi-guided rotor core ablation was performed to eliminate high PS sites, which achieved shorter duration of radiofrequency delivery to AF termination compared to a step-wise ablation [59, 60], but a similar AF-free maintenance in 12-month followups [61]. With a similar goal, the ablation of high DF regions was performed prior to the PV isolation [62] in 50 AF patients to eliminate high DF sites. The ablation of high DF sites was found to be associated with a higher post-ablation AF maintenance compared to PV isolation. However, another prospective ablation trial with 232 AF patients enrolled reported no such improvement [63].

Ablation of complex fractionated atrial electrograms (CFAEs) CFAEs refer to electrograms that are short CL, fractionated, and/or of low voltages over several cycles [64]. Regions with CFAE were believed to reflect sites of wave collision and slow conduction, and the anchor points of stationary reentries. Thus, ablation strategies targeting CFAEs were developed [65]. Mapping results by multi-polar catheters suggested the most common areas of CFAE are distributed on PV and their antra, coronary sinus, as well as LA roof and posterior wall [66], which are similar to where focal and reentrant drivers are distributed. However, as the strategy of PV isolation and CFAE ablation showed conflicting results in the follow-up studies between different centers, CFAE was considered as a passive phenomenon rather than an active driver for AF maintenance [67].

Voltage-guided and CMR-guided fibrosis ablation Myocardial fibrosis indicates an arrhythmogenic substrate, as it is responsible for wavebreaks, slow conduction, heterogeneity of APD and ERP. Areas with low voltage in the electrograms were shown to be related to myocardial fibrosis. Delayed enhancement magnetic resonance imaging (MRI) also provided possibilities in mapping fibrosis non-invasively. However, there was

no clear standard for the cut-off threshold in voltage mapping or pixel intensity in MRI images for localizing the fibrosis tissue [68]. An expert consensus recommended mapping and subsequently ablating low voltage areas only for non-paroxysmal AF patients undergoing repeated ablation [5].

2.2.4 Non-invasive AF mapping tools

The 12-lead electrocardiography (ECG) is the most commonly used clinical tool to diagnose the presence and mechanism of arrhythmia. As shown in Fig. 2.5, a 12-lead ECG system consists of three limb leads and six precordial leads, with all leads using Wilson's Center Terminal as a reference. These 12 leads then form the sagittal, frontal and transverse plane projections, and the morphologies of the ECG in these 12 leads can inform on electrical propagation in the heart. The potentials on leads of I, II, III aV_R, aV_L and aV_F in the 12-lead ECGs are obtained as follows:

$$I = V_L - V_R \tag{2.3}$$
$$II = V_R - V_F \tag{2.4}$$
$$III = V_L - V_F \tag{2.5}$$
$$aV_R = V_R - (V_L + V_F)/2 \tag{2.6}$$
$$aV_L = V_L - (V_R + V_F)/2 \tag{2.7}$$
$$aV_F = V_F - (V_R + V_L)/2 \tag{2.8}$$

The body surface potential maps (BSPMs) are similar to 12-lead ECGs, except with a larger number of electrodes, eg. 252 leads by the CardioInsight (Medtronic, NL) system in Fig. 2.6. BSPM provides a larger coverage of the potentials on the torso.

Electrocardiography imaging (ECGi) infers the epicardial maps of electrical potential by solving an inverse problem given the BSPMs and the geometries of the mesh and the torso acquired by computer tomography (CT) imaging. However, in order to cope with the ill-posed property of the inverse problem, ECGi uses regularization with prior assumptions. Successful demonstration of ECGi involves simple AF cases driven by a single focal or reentrant driver [41, 40]. However, so far, there is little validation using clinical data, especially for AF with multiple drivers [11]. Apart from this, ECGi also suffers from a generally low resolution, and can not distinguish between breakthroughs, spontaneous focal sources, as well as micro-reentries [70].

Other algorithms for non-invasive AF mapping using ECG and BSPMs were developed without the need for cardiac imaging. We introduced them as related works in Sec. 2.5.

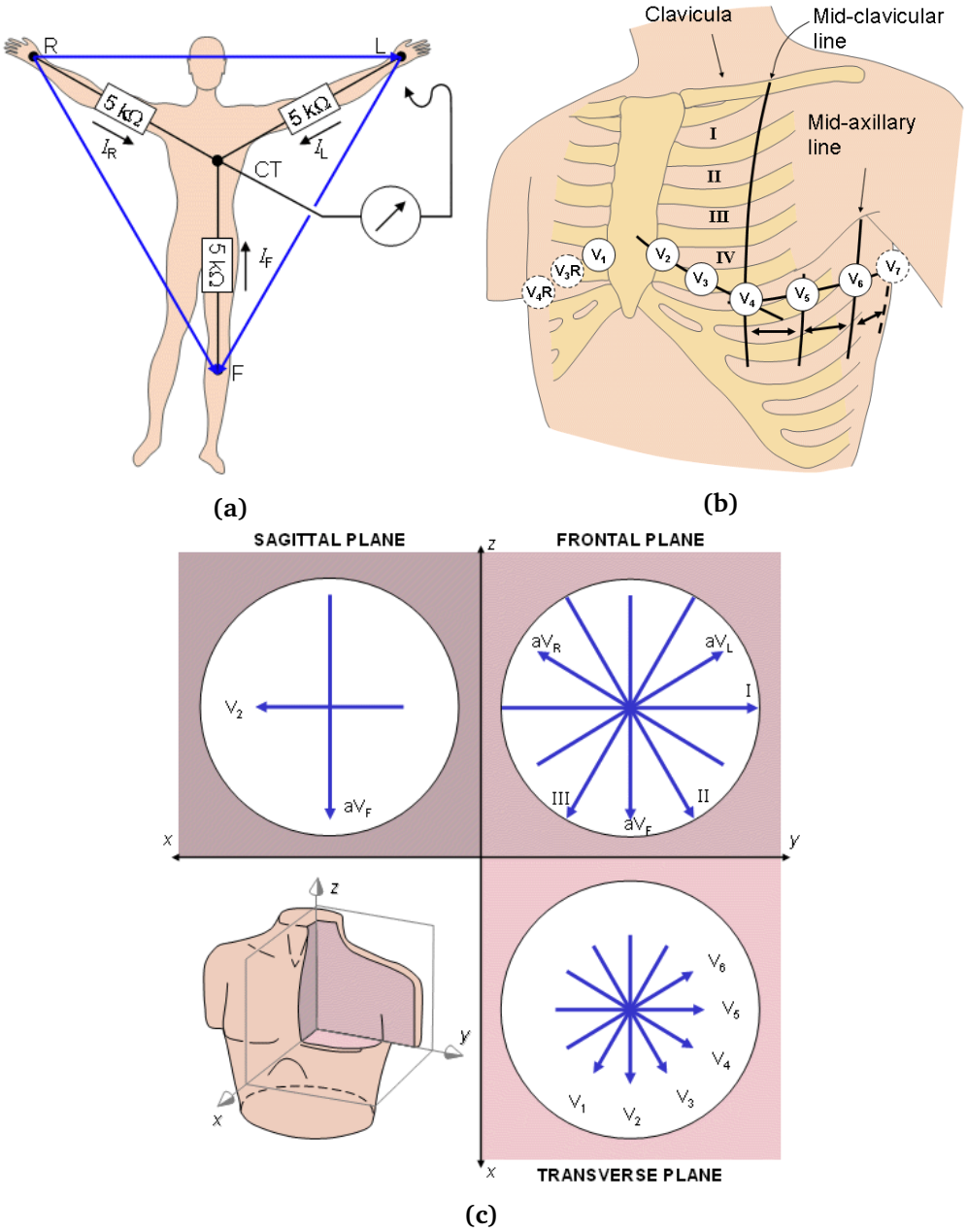


Fig. 2.5. (a) Lead placement of three limb leads (V_L, V_R, V_F)and Wilson center terminal (CT). (b) Lead placement of the six precordial leads in a 12-lead ECG system. (c) Interpretation of 12-lead ECG. Figure adapted from [69].

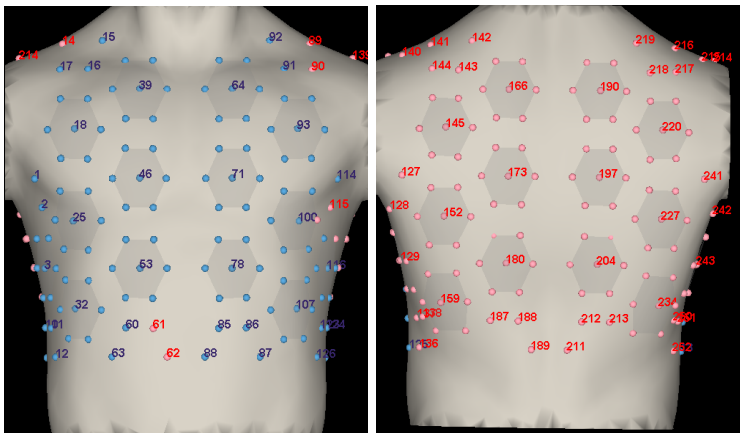


Fig. 2.6. Lead placement of BSPM in the CardioInsight system.

2.3 Basics of machine learning

Machine learning refers to computer algorithms that automatically identify patterns within data to accomplish tasks such as inference or prediction [16]. The ability for a machine-learning algorithm to correctly make predictions from the unseen data is called generalization which is valuable because the ultimate goal for machine learning is to make correct predictions on data from real-world applications.

If a model performs too well on the training dataset, but poorly on the unseen data, it is likely that there is an ***overfitting***. The opposite of overfitting is ***underfitting***, which means the model was unable to capture the relationship between the input and the output. In order to minimize the chances of overfitting and underfitting, when training a machine learning algorithm, a training dataset and a held-out test dataset are usually required. The training set is used to train the model while the test dataset remains invisible until evaluation. Sometimes the training set is further split into a training subset and a validation subset in order to select the hyperparameters of the model, where the model is trained on the training subset, and the evaluation on the validation dataset is used to select the best model.

Supervised machine learning refers to a class of data-driven methods with prediction targets available in the training/validation dataset, where the computer automatically extracts the statistical relationship between the inputs and the outputs. In a supervised classification problem, the outputs are discrete values denoting their classes, and they are available to the algorithm during the training phase. In our problem, the labels could be the cardiac arrhythmic conditions or mechanisms, or the outcome of a treatment. In this work, Gaussian processes and the classification models belong to supervised learning. ***Unsupervised machine learning*** refers to the problem where the training/validation dataset does not contain the prediction targets. In this work, principal component analysis, blind source separation problem, and clustering belong to unsupervised learning.

2.3.1 Gaussian Processes

A Gaussian process (GP) is a generalization of the Gaussian probability distribution over functions [71]. A GP is uniquely determined by its mean and covariance, which is also called the kernel function. The function f drawn from the GP is denoted as $f(\mathbf{x}) \sim \mathcal{N}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$ where

$$m(\mathbf{x}) = E\{f(\mathbf{x})\} \tag{2.9}$$
$$k(\mathbf{x}, \mathbf{x}') = E\{(f(\mathbf{x}) - m(\mathbf{x}))^T (f(\mathbf{x}') - m(\mathbf{x}'))\} \tag{2.10}$$

where $E\{\cdot\}$ denotes the expected value of the random variable enclosed. A GP models the function f as a Bayesian linear model with "kernel tricks". By mapping the input to the feature space by a function $\phi(x)$, the GP is able to model a variety of functions using a linear regression, as the mapping function can be nonlinear, and the feature space can be of any finite number of dimensions. To see why this is the case, consider the following regression model, where the mapping between the feature and the output is linear, with a noise term $\mathbf{w}$

$$f(\mathbf{x}) = \phi(\mathbf{x})^T \mathbf{w}, \text{ where } \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma_p) \tag{2.11}$$

Following Eq. (2.9) and (2.10), the mean $m(f(\mathbf{x}))$ and the covariance $k(f(\mathbf{x}), f(\mathbf{x}'))$ of the function $f(\mathbf{x}) \sim \mathcal{N}(m(\mathbf{x}), k(f(\mathbf{x}), f(\mathbf{x}')))$ are

$$m(f(\mathbf{x})) = E\{\phi(\mathbf{x})\mathbf{w}\} = \phi(\mathbf{x})E\{\mathbf{w}\} = \mathbf{0} \tag{2.12}$$
$$k(f(\mathbf{x}), f(\mathbf{x}')) = \phi(\mathbf{x})^T E\{\mathbf{w}\mathbf{w}^T\}\phi(\mathbf{x}') \tag{2.13}$$
$$= \phi(\mathbf{x})^T \Sigma_p \phi(\mathbf{x}') \tag{2.14}$$

where $\Sigma_p = E(\mathbf{w}\mathbf{w}^T)$. It can be seen that the output of the kernel function $k(f(\mathbf{x}), f(\mathbf{x}'))$ is an inner product of the features, $\phi(\mathbf{x})^T \Sigma_p^{1/2}$ and $\Sigma_p^{1/2}\phi(\mathbf{x})$.

GP prior and posterior Similar to other Bayesian updates, before seeing any observations, the GP models the prior distribution of the target function. When new observations are added, the posterior distribution of the target function can be obtained, which is what we are interested in.

Consider the observations $\mathbf{y}$ on the inputs $\mathbf{x}$ mixed with white noise ϵ, $\mathbf{y} = f(\mathbf{x}) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$. We used $(\mathbf{X}, \mathbf{y}) = \{(\mathbf{x}_i, y_i)\}$ to denote n pairs of training input vectors $\mathbf{x}$ and scalar ouputs (training targets) y, $\mathbf{X}$ is the training input matrix which aggregates $\mathbf{x}$ into n rows, and $\mathbf{y}$ is a vector of all scalar outputs. We also denoted a testing input matrix $\mathbf{X}^*$ for the test inputs.

Following Eq. (2.12) and (2.14), the prior distribution of the prediction on the test

inputs $f(\mathbf{X}^*) = \mathbf{X}^*$, $\mathbf{f}^*$ is given by

$$\mathbf{f}^* \sim \mathcal{N}(\mathbf{0}, \kappa(\mathbf{X}^*, \mathbf{X}^*)) \tag{2.15}$$

where $K(\cdot, \cdot)$ is the covariance matrix, which contains the mapping function $\phi(\mathbf{x})$ implicitly, as shown in Eq. (2.14). The samples drawn from the posterior distribution are shown in Fig 2.7 (a). The joint distribution of the observations $\mathbf{f}^*$ and the prediction of the target function $\mathbf{y}$ are given as follows:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}^* \end{bmatrix} \sim \mathcal{N}\Big(\mathbf{0}, \begin{bmatrix} \kappa(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I, \kappa(\mathbf{X}, \mathbf{X}^*) \\ \kappa(\mathbf{X}^*, \mathbf{X}), \kappa(\mathbf{X}^*, \mathbf{X}^*) \end{bmatrix}\Big) \tag{2.16}$$

where I is an identify matrix. After seeing a number of observations, the distribution of the predicted function $\mathbf{f}^*$, which is the posterior distribution, is given by

$$\mathbf{f}^* | \mathbf{X}^*, \mathbf{f}, \mathbf{X} \sim \mathcal{N}(m(\mathbf{f}^*), cov(\mathbf{f}^*)) \tag{2.17}$$

$$\text{where } m(\mathbf{f}^*) = \kappa(\mathbf{X}^*, \mathbf{X})(\kappa(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I)^{-1}\mathbf{y} \tag{2.18}$$

$$cov(\mathbf{f}^*) = \kappa(\mathbf{X}^*, \mathbf{X}^*) - \kappa(\mathbf{X}^*, \mathbf{X})(\kappa(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I)^{-1}\kappa(\mathbf{X}, \mathbf{X}^*) \tag{2.19}$$

The samples drawn from the posterior distribution are shown in Fig 2.7 (b). When there are observations taken nearby, the standard deviation of the posterior distribution is smaller and the estimated mean is closer to the true function.

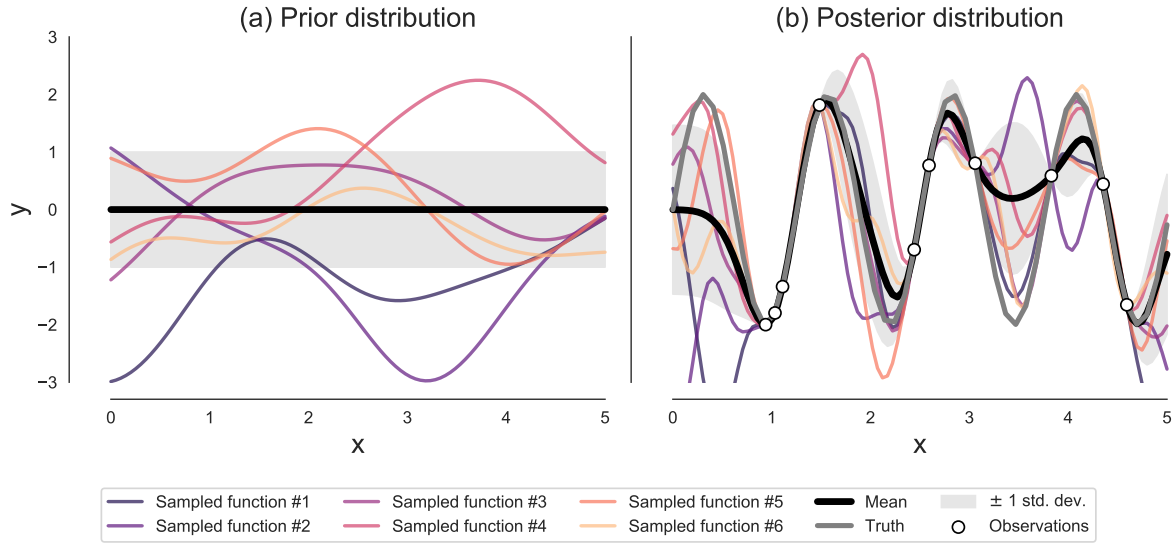


Fig. 2.7. The sample functions drawn from (a) prior distribution and (b) posterior distribution of GP functions, with their mean and covariance, over 10 noise-free observations of the true function $y = 2\sin(5x)$**.** The kernel function is a radial-basis function with one hyperparameter of characteristic length-scale l bounded to $0.1 \leq l \leq 10$. In (b), l was optimized by maximization of marginal likelihood, and the mean of the posterior distribution approximates the truth function. A better performance can also be observed at locations where the uncertainty bound (± 1 standard deviation (std. dev.)) is smaller as observations are taken nearby.

Optimization of GP kernel As can be seen in Eq. (2.19), it is essential to select a suitable kernel in order to estimate the target function. For example, in many covariance functions, there is usually a hyperparameter specifying the characteristic length-scale, which, loosely speaking, means how far apart two data points in the input space can be before no longer considered correlated. A very large value of characteristic length-scale gives rise to smoother GP estimated functions akin to "underfitting", whereas using a length-scale value that is too small may lead to "overfitting".

Two common methods for selecting kernel functions and their hyperparameters include maximization of marginal likelihood, and cross validation. For the former method, the marginal likelihood contains a data fit term, a complexity penalty and a constant term, which can be optimized by following the gradient of the marginal likelihood with respect to the hyperparameters θ. The marginal likelihood is given by marginalizing the prediction term $\mathbf{f}$ from the joint distribution in

$$p(\mathbf{y}|\mathbf{X}) = \int p(\mathbf{y}|\mathbf{f}, \mathbf{X}) p(\mathbf{f}|\mathbf{X}) \mathrm{d}\mathbf{f} \tag{2.20}$$

The marginal likelihood is therefore

$$\log p(\mathbf{y}|\mathbf{X}, \theta) = -\frac{1}{2}\mathbf{y}^T(\kappa + \sigma_n^2 I^{-1})\mathbf{y} - \frac{1}{2}\log|\kappa + \sigma_n^2 I| - \frac{n}{2}\log 2\pi \tag{2.21}$$

where $\kappa = \kappa(\mathbf{X}, \mathbf{X})$.

Notice that in Eq. (2.21), we explicitly write the hyperparameter variable θ as we can then take the derivative of $p(\mathbf{y}|\mathbf{X}, \theta)$ with respect to θ and perform gradient descent. As with other gradient descent methods to find the optima, starting from several different locations increases the likelihood of finding the global maximum of the marginal likelihood. For the cross-validation method, closed forms of the partial derivative of the cross-validation predictive mean and variances with respect to the hyperparameters are also derivable, which are used to update the hyperparameters. The formulations for these two updates can be found by Rasmussen and Williams [71].

GP for regression (GPR) GPR has been frequently applied in modeling unknown functions that can only be evaluated by a limited set of observations. Advantages of the Bayesian framework for performing regression tasks include the available uncertainty estimation, and the flexible description of the underlying data structure via the adoption of a kernel.

The GPR was previously used to represent the estimation and uncertainty of ventricular voltage maps from a finite set of CARTO measurements for patients of congenital heart diseases [72], which was used to guide the mapping process. A Gaussian Markov Random Field, a Bayesian regression method using a kernel similar to GPR, was

also used to estimate the local activation map of AF from CARTO measurements [73]. In addition, GPR was used to model data in the spectral domain containing periodic components, which entailed using periodic covariance functions and then fitting the parameter [74, 75], as well as a pure non-parametric model with application in estimating heart rate variability [76].

2.3.2 Blind Source Separation

Physiological measurements are usually collected by multiple sensors to record the activity of organs, such as the electrocardiogram and the magnetocardiogram for heart activities, and the magnetoencephalogram and the electroencephalogram for brain activities. These measurements are spatiotemporally correlated far-field signals. When a clinician inspects the signals, they look for spatiotemporal patterns in distinct activities, with a mental process equivalent to source separation [77].

The problem of estimating these sources without knowing the specifics of the transmission system is termed ***Blind Source Separation*** (BSS). It is surely not possible to estimate the sources, without having *a priori* knowledge of the source characteristics. It turns out however that one can estimate the sources up to certain indeterminacies, including arbitrary scaling, permutation and the delay of estimated source signals, while still preserving the temporal structure of the source waveforms [78].

BSS methods infer the sources of measurements based on *a priori* knowledge about the sources, without knowing the construction of transmission systems. This is different from the inverse problem of ECGi, which infers the sources by explicitly utilizing the geometries of the heart and torso, as well as the transformation matrix for forward modeling as inputs. BSS operates on signals after the preprocessing steps of centering, whitening, and band-pass filtering, as well as dimension reduction techniques such as the Principal Component Analysis, as introduced below.

Suppose signals are collected using N sensors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]^T$, and there are M real sources $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M]^T$. The mixing function f in $\mathbf{X} = f(\mathbf{S})$ project the $\mathbf{S}$ to the signals $\mathbf{X}$. If the transform f is linear, we can estimate a mixing matrix consisting of K orthogonal vectors $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_K]$. The unmixing matrix is the transpose of the mixing matrix, which projects the signal $\mathbf{X}$ to K sources, given by

$$\mathbf{X} = \mathbf{U}\mathbf{S}$$
$$\mathbf{S} = \mathbf{U}^T\mathbf{X}$$

Centering and Whitening

The centering transform, given by $\mathbf{X} - \bar{\mathbf{X}}$, and the whitening transform, are important preprocessing steps for BSS in order to remove the correlation between channels of signals, and for the sources to have unit variance. For the signal matrix after whitening operation $\mathbf{X}_w$, we have

$$E\{\mathbf{X}_w \mathbf{X}_w^T\} = \mathbf{I} \tag{2.22}$$

The whitening transformation also relies on the eigenvalue decomposition of the covariance

$$E\{\mathbf{X}\mathbf{X}^T\} = \mathbf{U}_w \mathbf{D}_w \mathbf{U}_w^T \tag{2.23}$$

where $\mathbf{U}_w$ is an orthogonal matrix of N eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N$, and $\mathbf{D}_w$ is a diagonal matrix with eigenvalues on the diagonal. The whitened signals are given by the following equation which satisfies Eq. (2.22).

$$\mathbf{X}_w = \mathbf{D}_w^{-1/2} \mathbf{U}_w^T \mathbf{X} \tag{2.24}$$

Principal Component Analysis

Data in real-world applications is usually complex and of high dimension, but usually lies close on a low-dimensional non-linear manifold spanned by latent variables [16]. The simplest continuous linear latent model is Principal Component Analysis (PCA) [16]. PCA is a widely used technique for dimension reduction, feature extraction and data visualization in various applications. It decomposes data into a set of orthogonal components by maximizing the variance of the data points on the orthogonal axes, which is equivalent to maximizing uncorrelatedness between the sources [77]. The mean and the variance of the sources are

$$E\{\mathbf{s}_1\} = E\{\mathbf{u}_1 \mathbf{x}\} = \mathbf{u}_1 \bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{u}_1 \mathbf{x_i} \tag{2.25}$$

$$Var\{\mathbf{s}_1\} = Var\{\mathbf{u}_1 \mathbf{x}\} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{u_1}^T \mathbf{x_i} - \mathbf{u_1}^T \bar{\mathbf{x}})^2 = \mathbf{u_1}^T \mathbf{S} \mathbf{u_1} \tag{2.26}$$

where $\mathbf{S} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x_i} - \bar{\mathbf{x}})(\mathbf{x_i} - \bar{\mathbf{x}})^T$. The transform vector $\mathbf{u_1}$, which maps $\mathbf{x}$ to the first component, can be obtained by maximizing the variance $Var\{\mathbf{s}_1\}$ subject to the constraint of $\mathbf{u_1}^T \mathbf{u_1} = 1$ using a Lagrange multiplier [16]:

$$\mathcal{L} = \mathbf{u_1}^T \mathbf{S} \mathbf{u_1} + \lambda_1 (1 - \mathbf{u_1}^T \mathbf{u_1}) \tag{2.27}$$

Setting $\frac{\partial \mathcal{L}}{\partial \mathbf{u_1}} = 0$, it follows that

$$\mathbf{S}\mathbf{u_1} = \lambda_1 \mathbf{u_1} \tag{2.28}$$

$$\lambda_1 = \mathbf{u_1}^T \mathbf{S}\mathbf{u_1} = Var(\mathbf{x}) \tag{2.29}$$

The maximal variance is then given by the maximal value of λ_1. All PCA projected components can be estimated in a similar manner, where all N projected components can be incrementally retrieved by using the eigenvalues of subsequent amplitudes $\lambda_1, \lambda_2, \ldots, \lambda_N$. In this manner, the percentage of using K PCA components to reconstruct the signals could be calculated as

$$\frac{\sum_{i=1}^{K} \mathbf{u_i}^T \mathbf{S}\mathbf{u_i}}{\sum_{j=1}^{N} \mathbf{u_j}^T \mathbf{S}\mathbf{u_j}} = \frac{\sum_{i=1}^{K} \lambda_i}{\sum_{j=1}^{N} \lambda_j} \tag{2.30}$$

Second-Order Blind Source Separation (SO-BSS)

A set of BSS algorithms are based on second-order statistics, known as SO-BSS methods. These algorithms maximize the source periodicity through calculating the "average eigen-strcture" of the covariance matrices [79]. In these methods, a time-delayed covariance matrix with a time lag of τ is defined as

$$C_x(\tau) = \mathbf{X}_{\tau+1:T} \mathbf{X}_{1:T-\tau}^T \tag{2.31}$$

where the subscripts of $\mathbf{X}_{i:j}$ denoting the times from i to j are included, and the subscripted T represents the length of the signal.

In SO-BSS, the transform f is assumed to be linear and stationary, and the transformation matrix U is estimated via eigenvalue decomposition, or singular value decomposition, which reveals the joint structures shared among the covariance matrices. This idea was initially proposed in the AMUSE algorithm [80] to deal with two such covariance matrices, one of which has zero time lag $\tau = 0$, so that a general eigenvalue decomposition can be used to diagonalize the two covariance matrices, $C_x(0)$ and $C_x(\tau)$. As the decomposition is sensitive to the choice of τ, the algorithm was later developed into a joint diagonalization process on a large set of covariance matrices corresponding to multiple time lags as presented in the SOBI algorithm [81]. The joint diagonalization provides the eigenvectors shared amongst different covariance matrices, where the eigenvalue matrices $\tilde{\mathbf{D}}_0, \tilde{\mathbf{D}}_1, \ldots, \tilde{\mathbf{D}}_K$ may be different but the eigenvectors in the matrices $\tilde{\mathbf{U}}$ are shared.

$$\tilde{\mathbf{U}} \tilde{\mathbf{D}}_i \tilde{\mathbf{U}}^T = C_x(\tau_i), i = 1, 2, \ldots, K$$

Combined with the whitening process in Eq. (2.24), the sources are given by

$$\tilde{\mathbf{S}} = \tilde{\mathbf{U}}^T \mathbf{X}_w$$

One of the biggest advantages of the SO-BSS is that since they extract sources according to the eigenvalues from the set of covariance matrices, these algorithms produce estimated sources in a unique and consistent ordering [79]. In many applications, the number of sources is unknown, such as the number of rotors in an AF episode. SO-BSS exempted the need to estimate the source number, which is usually required in other types of BSS problems.

Independent component analysis (ICA)

ICA is a special form of BSS problem, which infers the sources by maximizing the ***independence*** or ***decorrelatedness*** between sources. A class of approaches is to maximize the non-Gaussianity measure between the estimated components. The first iteration of a well-known ICA algorithm, the FastICA algorithm [82], maximizes the fourth moment of the sample data, i.e. the kurtosis measure, in order to minimize the non-Gaussianity. It was later improved by an approximation of negentropy [83] to overcome the sensitivity to outliers for estimation of non-Gaussianity by the kurtosis, which is equivalent to minimizing the mutual information between the sources.

Another commonly deployed ICA method, InfoMax [84], maximizes the entropy of the components via a neural network. It was proved later that the InfoMax approach is equivalent to maximum likelihood estimation of the sources, and given that the density of the sources is accurate, InfoMax is equivalent to minimizing the mutual information [85]. These methods can estimate a non-linear transform function f, but there is no ordering between the sources, so prior knowledge about the number of sources is required.

2.3.3 Random Forest Classifier

One of the most widely used nonlinear classifiers is the random forest model [86], which is an ensemble method based on decision trees. The decision tree builds decision rules in a tree-like structure. To train a decision tree, in each node, an attribute (as well as a binary cut-off threshold if the variable is a continuous variable) is selected to split the data arriving at the current node into positive and negative examples for the following nodes. The attribute is determined by the one with the maximal information gain of the dataset S the node sees, measured by the entropy as follows

$$Entropy(S) = -p_+ \log_2 p_+ - p_- \log_2 p_- \tag{2.32}$$

where p_+ and p_- denote the proportion of the positive and negative examples for using selected attribute and threshold to split S, respectively. The decision tree greedily optimizes such measures until a certain depth, usually specified by the user, or until all attributes are exhausted.

A random forest is an ensemble of decisions trees, which uses bootstrapping methods to sample subsets of data with placement. A decision tree is built for each subset, and the final classification result uses the majority vote of the trees, thereby reducing the usual overfitting compared to using one decision tree [86].

2.3.4 Deep Learning

With the maturity of deep learning methods, the interpretation of ECG signals has been diverted to deep learning methods. Here, we describe the mathematical background for common deep learning practices. Deep learning is a special kind of machine learning technique, which uses multiple connected layers of neurons to approximate the mapping between the input and the target output data. Similar to other optimization problems, a loss function that measures the distance between the output and the target is optimized, usually by updating the weights of a neural network by a ***backpropagation*** process [87]. During the backpropagation, the network parameters are updated according to the partial derivative between the loss with respect to each parameter of each layer backwards (i.e. from the last to the first layer), calculated using the chain rule. A stochastic gradient descent method is usually used for optimization, which performs the gradient descent on small batches of data, to improve performance and reduce memory consumption compared to gradient descent on the whole dataset.

Basic architecture deep neural network

A basic deep learning architecture is a feedforward neural network with multiple layers, which is also called a deep ***feedforward neural network*** or a ***multilayer perception*** (MLP).

Output layer The last layer of the neural network is called the output layer. For classification tasks, the output layer is often chosen to be a sigmoid layer, which transforms the inputs of the layer $\mathbf{h}$ into probalistic outputs $p \in [0, 1]$.

$$\hat{y} = \sigma(\mathbf{w}^T \mathbf{h} + \mathbf{b}) \tag{2.33}$$

where $\mathbf{w}$ is the weight, and $\mathbf{b}$ is the bias.

A binary cross-entropy loss between the probabilistic output vector $\mathbf{p}_i$ and the binary ground-truth vector $\mathbf{y}_i$ for the i-th class, over N_C classes, is often used as an optimization goal for classification. A basic form of the loss is defined as

$$\mathcal{L} = -\sum_{i=1}^{N_C} [\mathbf{y}_i \cdot \log \mathbf{p}_i + (1 - \mathbf{y}_i) \cdot \log(1 - \mathbf{p}_i)]$$

where the $\cdot$ denotes the dot product of two vectors.

Hidden layers The layers between the input layer, the first layer, and output layer are called hidden layers. While hidden layers can be any function, such as a fully connected function,

$$g(z) = \mathbf{w}^T \mathbf{h} + \mathbf{b} \tag{2.34}$$

in order learn more complex features, a non-linear transformation function is needed. A popular choice for the non-linear activation function in modern deep learning networks is to use rectified linear unit (ReLU) [88], defined as

$$g(z) = \max(0, z)$$

Furthermore, different distributions of input data faced in each layer may cause an "internal covariate shift" problem where each layer has to be adapted to the different distribution. This could result in a significant shift due to the cascading effect of the backpropagation in a deep neural network. To solve this, a ***batch normalization*** transform [89] to normalize the input data of each layer is usually added to the neural network as hidden layers.

Convolutional neural network (CNN)

Generally speaking, increasing the number of layers and the number of parameters in each layer will increase the flexibility of the model. However, increasing network layers introduces the problem of "vanishing gradients" with an inefficient weight update. The increased number of parameters is also likely to cause overfitting, just as with other machine learning models with too many parameters. To solve this problem, CNN and recurrent neural network have been developed, which can learn local features efficiently via parameter sharing.

CNNs have been successfully applied to learning invariant features in many applications with a regular grid structure, including regularly-sampled time series as 1D grids of values, and image and video data as 2D and 3D grids of pixels [17]. A CNN is

a deep neural network that uses a convolution operation in at least one of the layers, which is often combined with pooling layers.

Convolution layer The convolution layer performs a standard feature extraction operation, with parameters w updated during the training. The formulas for 2D convolution operation are represented as

$$S(i,j) = (I * K)(i,j) = \sum_i I(i-m, j-n)K(m,n)$$
$$\text{or } S(i,j) = (I * K)(i,j) = \sum_i I(i+m, j+n)K(m,n)$$

Advantages of convolution are summarized as ***sparse interactions***, ***parameter sharing*** and ***equivariance to representation*** [17]. The sparse interaction is realized by using kernels much smaller than the input layer, which is typically between two fully connected layers. The same parameters are shared between different locations, which improves the generalization. The equivariant representation includes shift-invariance, meaning that the translation in the input space would produce the same change in the output space, which means, shifting the pixels in space does not impede object detection. It also enables learning long-range global features via multiple levels of composition through stacking multiple CNN layers.

Pooling layer The pooling operation is another important operation in a CNN. It takes the output of a pooling function from a rectangular region, to guarantee that the feature is invariant to small changes in the space. Max Pooling [90] is a standard choice, which selects a maximum value from a rectangular neighborhood. The pooling makes the representation in the next layer approximately invariant to small shifts in the last layer in regards to the presence of a feature [17].

Recurrent neural network (RNN)

RNNs were designed to learn long time-series. A time series of T samples

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$$

is fed as input through a basic RNN with the cells as shown in Fig. 2.8 (a-b). A cell at time t in the RNN architecture takes two inputs, the output from the previous cell $\mathbf{a}_{t-1}$, as well as the input data $\mathbf{x}_t$. These two vectors are multiplied by learn-able parameters and summed together, and transformed by a non-linear function, g_1, to produce the cell output $\mathbf{a}_t$. In the simplest case, g_1 can be the $\tanh$ function. The cell output can also be transformed by another activation function, g_2, such as a $softmax$ function for

classification tasks. The training labels can be supplied either as a sequential data, $\mathbf{y} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_T\}$, or a single entry as $\mathbf{y} = \hat{y}_T$.

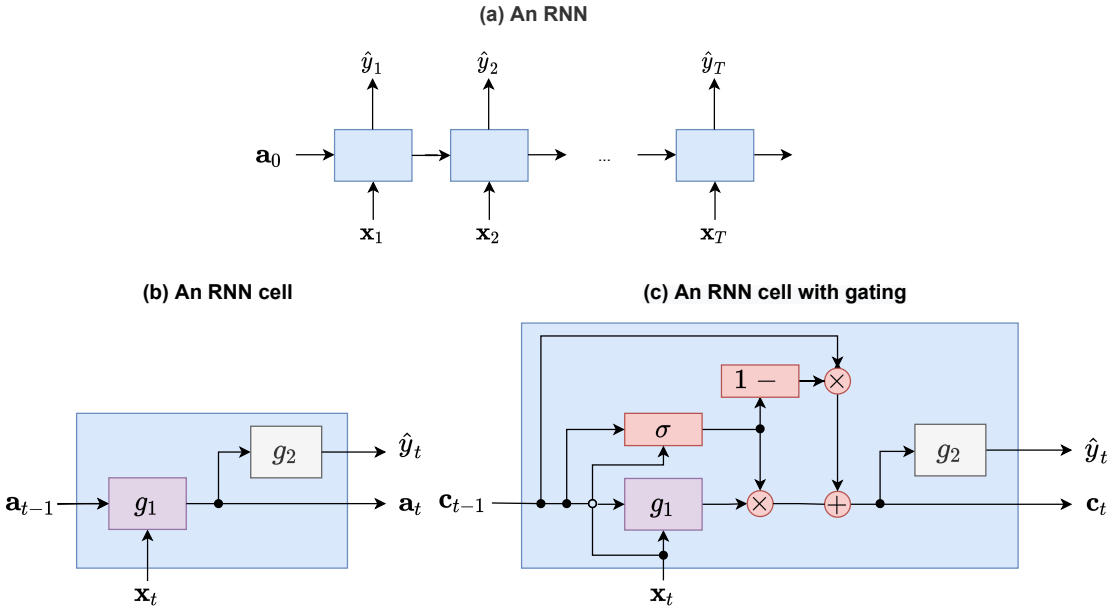


Fig. 2.8. The basic architecture of (a) an RNN, (b) an example RNN cell and (c) an example RNN cell with gating functions (in red).

The update rules of RNN cells can be described as:

$$\mathbf{a}_0 = \mathbf{0} \tag{2.35}$$

$$\mathbf{a}_t = g_1(W_a\,(\mathbf{a}_{t-1}, \mathbf{x}_t) + \mathbf{b}_a) \tag{2.36}$$

$$\hat{y}_t = g_2(W_y\,\mathbf{a}_t + \mathbf{b}_y) \tag{2.37}$$

where $\mathbf{a}_0$ denotes the input to the first layer has entries being all 0. The variables $\mathbf{b}_a$, $\mathbf{b}_y$, W_a and W_y are the biases and weight matrices in the linear layer before applying the functions g_1 and g_2, and $(\mathbf{a}_{t-1}, \mathbf{x}_t)$ denotes concatenation of the $\mathbf{a}_{t-1}$ and $\mathbf{x}_t$.

However, training such an RNN on long sequences of time series still suffers from vanishing or exploding gradients. This problem is mitigated by adding gating variables to the cell, to improve its ability in processing long sequences. Two classical cell architectures are the Gated Recurrent Unit [91], and the Long Short Term Memory (LSTM) [92]. A basic RNN cell with gating function (Fig. 2.8 (c)) can be made by replacing the activation output $\mathbf{a}_t$ above by a memory-enabled output $\mathbf{c}_t$. An update gate variable, Γ_u, decides whether the output of a time step t, $\mathbf{c}_t$ is updated by a candidate $\tilde{\mathbf{c}}_t$ or the previous input. Mathematically, the update equations in Eq. (2.36) are replaced by following equations:

$$\tilde{\mathbf{c}}_t = g_1(W_c\,(\mathbf{c}_{t-1}, \mathbf{x}_t) + \mathbf{b}_c) \tag{2.38}$$

$$\Gamma_u = \sigma(W_u\,(\mathbf{c}_{t-1}, \mathbf{x}_t) + \mathbf{b}_u) \tag{2.39}$$

$$\mathbf{c}_t = \Gamma_u \, \tilde{\mathbf{c}}_t + (1 - \Gamma_u) \, \mathbf{c}_{t-1} \tag{2.40}$$

where $\mathbf{b}_c$, $\mathbf{b}_u$, W_c and W_u are the biases and weight matrices in the linear layer before applying the function g_1 and the sigma function σ, and $(\mathbf{c}_{t-1}, \mathbf{x}_t)$ denotes concatenation of the $\mathbf{c}_{t-1}$ and $\mathbf{x}_t$.

The candidate $\tilde{\mathbf{c}}_t$ is the output of a sigmoid function σ over a linear function of $\mathbf{c}_{t-1}$ and $\mathbf{x}_t$. When the update gate variable Γ_u is equal to 0, the previous input $\mathbf{c}_{t-1}$ is taken as $\mathbf{c}_t = \mathbf{c}_{t-1}$, thereby, being preserved to the next time. On the contrary, when the Γ_u is equal to 1, the candidate $\tilde{c}_t$ is adopted to be the next output, $\mathbf{c}_t = \tilde{\mathbf{c}}_t$.

To further enhance the capability of RNNs, they can be trained in two directions, forwards and backwards, so that the information can flow from an earlier time to a later time, as well as vice versa. Stacking more than one RNN cells for each time step encodes richer latent information for each time step, which can be suitable to learn data with more complex structures [17].

Graph neural network (GNN)

A point cloud is a common data representation, especially for many applications in computer vision such as shape classification, object detection, point cloud segmentation. The point cloud representation is also suitable for describing sensor placement for physiological measurements, as the sensors are rarely placed as a regular grid. Deep learning on point cloud data has been studied for decades. Albeit powerful, the CNN and RNN described above are only suitable for regular grid-like data. Earlier attempts to deal with the point cloud used convolution operations, by representing the point cloud as sparse volumetric data so that 3D convolution could be applied [93]. As point clouds usually arise from computer vision applications where cameras were used, multi-view approaches were also popular [93]. A point-wise MLP method, PointNet [94], was also proposed to learn the 3D classification by explicitly extracting the point-wise features of 3D points through multi-layer MLPs [93]. However, PointNet was unable to capture local structures essential in the extraction of global features [95].

The high demand across different applications has given rise to considerable research attention into deep learning techniques on non-Euclidean data, such as a deep neural network operating on a graph, which forms the field of ***geometric deep learning*** [96]. A simple and elegant idea to mimic the convolution operation on a regular grid to irregularly placed graph nodes is via ***message passing*** between nodes [97]. The message passing employs the following operations to pass updates from neighbouring nodes $\mathbf{x}_j$ of each node $\mathbf{x}_i$, to the node $\mathbf{x}_i$ [98]

$$\mathbf{x_i} = \gamma(\mathbf{x}_i, \square_{j \in \mathcal{N}(i)} \phi(\mathbf{x}_i, \mathbf{x}_j, e_{j,i})) \tag{2.41}$$

where ϕ denotes a differentiable function, such as MLP, that operates on both nodes and the edge $e_{j,i}$ which propagates the message from $\mathbf{x}_j$ to $\mathbf{x}_i$. The $\square$ denotes a differentiable, permutation-invariant function, such as sum, mean or max, for assembling all the messages from every neighbour of i, $\mathcal{N}(i)$. Another differentiable function γ then processes this propagation message and assembles it with the $\mathbf{x}_i$. This update method can be accelerated on Graphical Processing Units [98], and can act as a template for many state-of-the-art GNN architectures such as an edge convolution operation [99].

2.4 Multiscale modeling for AF

2.4.1 Construction of computer models

Multiscale modeling integrates multiple scales of models, such as cell, tissue, organ and body. We provide a brief summary of the milestones in building multi-scale models for AF. In 1964, Moe et al. [44] developed a first tissue-level AF model using a cellular automata model, on which they demonstrated the multiple wavelets as the AF mechanism. The first three-dimensional (3D) AF model was developed by Harrild and Henriquez [100], which used a cellular model of Nygren et al. [26] with a realistic human atrial geometry and fiber, and inter-atrial connection structures including fossa ovalis, coronary sinus and Bachmann's bundle. They were able to reproduce the conduction pattern from several pacing sites of the atria observed in experiments. Vigmond et al. developed a morphologically realistic 3D atrial model, with the addition of discrete interatrial communication, anisotropic conduction, and muscular structures, to study the spatial distribution of reentries due to regional difference of electrical properties [34], which was later augmented by ACh regulation to increase APD heterogeneity [101].

Advancements in imaging technologies have enabled the construction of patient-specific 3D atria models with atrial geometry being represented by the following: (1) a 3D surface representation, containing a single manifold of atria, with geometries obtained from CT [102], or from invasively acquired electroanatomical maps [103]; (2) a 3D volumetric representation where geometries are represented as six-sided cubes or tetrahedra, with geometries acquired from MRI [104] or CT [105]; and (3) a 3D bi-layer representation, developed by Labarthe et al. [106, 107].

Due to the complexity in the atrial fiber structures, early modeling works [100, 34] assigned fiber structures manually with references to histology data or based on rules, which were improved by more detailed fiber structures obtained from histology [35]. In order to assign more realistic fibre structures obtained from detailed imaging data of ex-vivo hearts to patient-specific meshes, semi-automated rule-based methods and atlas-based methods [108] were developed. Recent works on atrial fibre structures have focused on fully automated methods to map fibres from one atrial mesh to others,

using predefined landmarks [105], or via a Universal Coordinate System shared across different atrial meshes [109].

The fibrotic structures alter wave propagation in the atria [110], constituting an important arrhythmogenic substrate. Atrial fibrosis is detected from LGE-MRI by means of reduced gadolinium clearance, and has been included in patient-specific AF models [108, 111]. Different modeling techniques of fibrosis including lower conductivity, edge splitting, and percolation were investigated by Roney et al. [112], and found that percolation was closest to the clinical values of the phase singularities extracted by ECGi. Nonetheless, further validation for the modeling methodologies of atrial fibrosis using clinical data is still needed [110].

2.4.2 Computer models aided mechanistic understanding

There has been a long history of using 3D atria models incorporating multi-scale electrophysiological properties to understand the mechanisms that induce or sustain AF. For example, Jacquemet et al. [113] used a monolayer model to validate that the value of wavelength, which is the product of ERP and conduction velocity, should be under a certain threshold for sustained reentries, and a shortening of wavelength leads to an increase of AF vulnerability. The same type of monolayer model was later used to show an inverse relationship between AF CL and AF vulnerability in Hassaiguerre et al. [114]. On the 3D volumetric model of Harrild and Henriquez [100], Gong et al. [115] studied the repolarization dispersion as the mechanism underlying repetitive ectopic beats in paroxysmal AF. On three bilayer atrial meshes built from patient MRI images, Roney et al. studied the effects of PV in initiating reentries from PV triggers include APD shortening and conductivity slowing [107], and on the same meshes, Saha and Roney et al. studied how the presence of fibrosis affects the locations of phase singularities [116].

2.4.3 Computer models optimized AF ablation strategies

Ablation strategies have been optimized by inducing AF and performing virtual ablation on the 3D atria model. On a general bilayer biatrial model with atrial fibrosis, Bayer et al. [117] investigated several different ablation strategies, including PV isolation, PS ablation, and a line ablation streamlining the wave propagation to resemble sinus rhythm with multiple lines, and concluded that the last line ablation strategy achieved the highest efficacy. On a prospective study of catheter ablation for 10 persistent AF patients with fibrosis in the atria, Boyle et al. [118] built personalized 3D atria models to predict the optimal ablation sites for persistent AF, which resulted in 90% patients remaining AF-free in an average of 309 post-ablation days. Loewe et al. [119] simulated depolarization wave propagation on MRI-derived 3D atria models, and extracted all conduction

pathways as vulnerability paths to be avoided, in order to reduce the development of atrial flutter following an AF ablation therapy. They demonstrated retrospectively the presence of such atrial flutter vulnerability paths from a patient developing an atrial flutter after an AF ablation. Shade et al. [120] used features derived from LGE-MRI images and AF simulations by fast-pacing on 30 sites uniformly on the LA. Using machine learning extracted from simulated episodes, including the number of rotors and macro-reentries, they predicted the AF recurrence post-PV isolation, which achieved validation sensitivity and specificity scores of 82% and 89% with ten-fold nested cross-validation on 32 patients.

2.5 Related work on analysis of BSPM/ECG signals for AF treatment

In the last section of the chapter, we review state-of-the-art methods for interpreting BSPM/ECG signals for AF prediction and treatment. We summarize the related works in Table. 2.1, which are divided into time-domain, frequency-domain and equivalent source-domain methods. As can be seen, the interpretation of ECG/BSPM for AF mapping has also been aided with forward modeling of ECG/BSPMs by multiscale modeling. Further comments on the papers are below.

2.5.1 Time-domain analysis

Atrial activity organization

Complexity is a physiological property arising from information theory, which measures the randomness of a process. For example, the process of tossing a fair dice contains the maximal entropy, as it is sufficiently hard to guess which face it shows. A widely used complexity measure for ECG signals is the sample entropy [144], which was shown to be increased within the first minutes after AF onset and decreased before AF termination in AF patients [121]. It was also used to separate sinus rhythm with AF episodes [122]. Entropy measures extracted from recurrence quantification analysis of the first three PCA components of 12-lead ECGs were also developed, along with other features of spectral analysis, to distinguish between PV and non-PV drivers in ECGs of AF in order to predict acute PV isolation success [124].

Another complexity measure is related to the orthogonal bases of the signals from PCA analysis (Section 2.3.2). This includes “non-dipole index” measure, which takes the value of the total ratios of the unexplained variance of the first three PCA components (Eq. (2.30)) which was used to predict the immediate success of catheter ablation on

Table 2.1. Summary of related works. PersAF: Persistent AF. ParoxAF: Paroxysmal AF. AT: Atrial tachycardia. AFlut: Atrial flutter. VCG: Vectorcardiogrammes. Termination: acute termination following catheter ablation. Outcome: late AF recurrence after catheter ablation. MRAT: Macro-reentrant atrial tachycardia. BB: Bachmann's bundle. FO: fossa ovalis. CS: coronary sinus. PCA: Principal Component Analysis. AV: atrial-ventricular.

Domain	Category	Method	Reference(s)	AF pathophysiology	AF treatment	Data used
Time	Temporal	Sample Entropy	[121, 122]	Measure organization	Predict AF onset in ParoxAF patients [121]; Predict AF outcome at discharge [122]	45 ParoxAF patients [121]; 30 AF patients [122]
		PCA explained variance	[123, 15]	Measure organization	Predict AF termination after PersAF ablation	28 [123] and 97 [15] PersAF patients
		Recurrence quantification analysis	[124]	Discriminate PV and non-PV driver	Predict AF termination after ParoxAF ablation	Simulation on 16 atrial-torso geometries with 46 ParoxAF patients
		Maximal or average amplitude of f-wave	[14, 122]	Measure organization	Predict 1-year recurrence after PersAF ablation [14]; Predict AF termination after ablation [122]	90 PersAF patients [14]; 30 AF patients [122]
		Variance of of P-wave	[125]	Measure organization	Predict AF presence from Sinus Rhythm	73 PersAF patients and 20 control subjects
	Searching	BSPM clustering	[126, 127, 128]	Localize local site	Pre-operative planning	Simulation on single heart-torso geometry
	Spatiotemporal	Diopole trajectory	[129]	Measure organization	Pre-operative planning	Simulation on one heart-torso geometry
			[130, 131]	Localize focal site, distinguish rotor and FS	Pre-operative planning	Simulation on one heart-torso geometry
			[132]	Measure organization	Pre-operative planning	10 control and 10 ParoxAF patients
		VCG trajectory	[133]	Measure organization	Pre-operative planning	22 AF patients subject to AV junction ablation
			[134]	Identify MRAT subtypes	Pre-operative planning	Geometry-free simulations and 30 MRAT patients data
		Convolution neural network	[135]	Localize (undefined) driver	Pre-operative planning	Simulation on one heart-torso geometry
Frequency	Temporal	Organization index	[13, 136, 124]	Measure organization; distinguish PV and non-PV driver	Optimize timing of AF defibrillation; Predict AF termination after ParoxAF ablation [124]	Experimental data [13]; Simulation on 16 atrial-torso geometries with 46 ParoxAF patients [124]
		Spectral entropy	[137, 124]	Distinguish PV and non-PV driver	Predict AF termination after ParoxAF ablation [124]	Simulation on 16 atrial-torso geometries with 46 ParoxAF patients [124]
	Spatiotemporal	Highest DF	[12, 122]	Estimate highest DF sites [12]; Estimate dominant DF [122]	Pre-operative planning [12]; Predict AF outcome at discharge [122]	14 [12] and 30 [122] ParoxAF and PersAF patients
		PS trajectory	[138]	Localize AFlut loop	Pre-operative planning	9 AFlut patients
			[139]	Distinguish AT, AFlut and AF	Pre-operative planning; Predict AFlut termination	Simulation on one heart-torso geometry
Source	Spatiotemporal	BSS and/or ICA	[140, 141, 142, 143]	N/A	N/A	Various numbers of patient data

persistent AF [123, 15]. The separation of sources was also used in source-domain analysis discussed later (Section 2.5.3).

A higher amplitude in the lead V1 and the lead II in f-waves was also found to be related to younger patients with a shorter history of AF, and better treatment outcome for two-year follow-ups for in persistent AF patients [14]. Hernandez et al. also showed that the amplitude of P-waves can distinguish between sinus rhythm and an AF patient [122]. Therefore, we classified this feature into atrial activity organization. As a similar metric, higher variability of P-waves during sinus rhythm was observed in AF patients compared to control patients [125].

Spatiotemporal analysis using dipole trajectories

BSPM signals can be modeled by a fixed dipole model [69], where the electrical field is represented by a vector with three variables on a fixed position. The tip of the vector is represented in a 3D Cartesian or cylindrical coordinate system, and this gives rise to vectorelectrocardiogram (VCG). The mapping from BSPM to any 3D coordinate system is not unique, and entails information loss. Nonetheless, this mapping reduces the dimensions and is easier to process and provide diagnostic values.

Jacquemet [129] showed that the dipole trajectory extracted from simulated BSPMs showed an increasing complexity from sinus rhythm, to atrial flutter to AF. Giacopelli et al. [132] also showed that the dipole trajectory extracted from real BSPMs can separate healthy subjects and paroxysmal AF patients. Using simulated data, the dipole trajectories were further used to determine the focal sites on a set of patient geometries [130, 131]. Ng et al. [133] showed that the VCG has preferential routes from 22 AF patients, and the routes were correlated with atrial activity organization. Ruiperez-Campillo showed that the most common macroreentrant types, including the right (peritricuspid) and left (perimitral) macroreentrant circuits, can be inferred by the VCG [134].

The body surface Laplacian mapping was also computed from P-waves of healthy subjects, but was not applied on AF patients [145, 146]. As they reflected the source and the sink of the electrical field on the torso, they could potentially contribute to the diagnosis of AF mechanisms.

Searching by similarity

Computer models were also used to generate synthetic sets of BSPMs. The algorithm matched the nearest neighbour from the previously generated BSPMs, and so, no feature was explicitly defined for the signals. Several in-silico studies showed that this could be used for focal site prediction, including Potyagaylo et al. [126] and Ferrer-Albero et al. [127]. Godoy et al. [128] showed that this method would suffer from an

increased level of atrial fibrosis. However, these methods were trained and tested on the same set of atrial-torso geometry, which made any search-based solution trivial.

Deep learning

Deep learning techniques have also been applied to the analysis, but most of the methods were trained to separate AF from sinus rhythm, thanks to the abundance of ECG data from AF patients and control subjects. For AF treatment purpose, a recent published work [135] estimated drivers from BSPMs based on simulated data using a CNN (Section 2.3.4). Unfortunately, the definition of AF driver, which was an important concept of AF mechanism, was not clear to us from the manuscript. Furthermore, the study only used a single set of atrial-torso geometry.

2.5.2 Frequency-domain analysis

Spectral analysis using the Fast Fourier Transform transforms a time series from the time domain into the frequency domain. It was frequently used to analyze the cardiac signals including ECGs, in order to estimate the activation cycle length (CL) robustly.

Atrial activity organization

The organization index [13], which measures the content of the dominant frequency (DF) and its harmonics with respect to all frequency content, was a metric to measure atrial activity organization in the spectral domain. The organization index was used to optimize the timing of atrial defibrillation in animal studies [148]. The spectral entropy [137] was another measure of organization in the frequency domain, and although not directly developed for analyzing ECG initially, it was used as one of the candidate features to predict PV/non-PV driver [124].

Dominant frequency (DF) and phase singularity (PS)

The dominant frequency (DF) of a signal represents the frequency component with the highest energy. Guillem et al. [12] established that the highest DF region on the BSPMs corresponds well with the highest DF site in the nearest atrium, from 67-lead BSPM signals of 14 AF patients. The localization of high DF areas could help to narrow down the search scope for the highest DF during ablation. A limitation in the DF analysis is that it is agnostic to FS and reentrant sources. For complex AF mythologies, the frequency power spectrum is diffuse, making it hard to extract DF automatically.

The PS extracted in the phase map of the BSPMs is another important feature in spectral analysis. Using simulated data, Rodrigo et al. showed that the sustained PSs on the BSPM indicated the presence of rotors after high-pass filtering [149]. Liberos et al. [138] predicted the location and the chirality of atrial flutter by tracking the trajectory of PS. Marques et al. [139] demonstrated a potential use of PS trajectory in separating atrial tachycardia, atrial flutter and AF on 19 simulated BSPMs of one set of geometry. For PS analysis, attention should be paid to the fact that the PS was not a unique feature for reentrant sources [149]. There is also a higher systematic error between DF extracted on torso and in the heart for less regular atrial activities [136]. Therefore, it is unclear whether PS trajectory remains meaningful for rotor detection in complex AF episodes with multiple simultaneous wavelets. This is not surprising for far-field signals, and could be a theoretical limitation for how much we can utilize them for analyzing AF mechanisms.

2.5.3 Equivalent source-domain analysis

The BSPM is the result of mixing several atrial activities. To identify these atrial activities, methods to extract equivalent source separation were investigated, such as BSS methods described in Section 2.3.2. Using SO-BSS and/or ICA techniques, atrial content, opposed to the ventricular content, were extracted from BSPMs of AF patients [140, 141, 142, 143]. These methods were mainly used as a preprocessing step of signals for downstream tasks such as DF analysis.

2.5.4 Remarks on related works

From the previous works, it is evident that the patient follow-up data are the most direct indicator of therapeutic outcomes. Simultaneous mapping data of the atrium and the torso is also of great value to confirm AF mechanisms. However, it takes a long time and a good practice of data management to collect these types of data, so it is not always possible to obtain enough patient data that can be split into training, validation and testing sets. A robust validation further requires diagnostic cohort design, multi-center data collection and prospective deployment of the algorithm [150].

Computational modeling provides another data source for training and testing models, through the generation of synthetic data representative of AF patients. However, there are two observed problems from most of the previous methods on analyzing AF BSPMs that are trained on the simulated data. The first problem is that the current simulated methods oversimplified the AF episodes as driven by only one focal or reentrant driver [151]. For example, the prediction of the focal site was only based on the BSPMs of AF episodes where FS induced 1:1 activities to the rest of the atria [130,

131], which was rarely the case for FS in sustained AF episodes for AF patients. Non-invasive analysis of AF with a mixture of drivers is considerably harder. An example is that the prediction of focal site was hindered by a higher degree of fibrosis with altered conduction pathways, and subsequently, affecting the analysis of the BSPMs [128].

Another major problem is that many works (see data in Table 2.1) performed cross-validation testing on only one patient atrial mesh with one electrode set of BSPMs. Such designs ignored the fact that patients have different atrial and torso geometries. Using computer modeling, we learned that the morphologies of ECGs are impacted by the site of activation, including the sino-atrial node [152] as well as geometry, especially changes such as LA dilation [153], suggesting the importance of considering different sets of atria-torso geometries. Although methods trained only on one set of heart-torso geometries are usually proof-of-concept, it should go through additional validation before being applied to patient data. For a new patient, the only way to apply methods that have been trained and tested using the same geometry is to generate computer simulations using personalized geometries, and to retrain machine learning models on the patient-specific simulations. Such operations, however, are rather expensive to deploy in clinical practice.

We conclude that there is still a high demand to develop machine learning algorithms using a large and diverse set of training data containing realistic BSPMs generated from computer simulations, to develop the generality for the diverse patient population. It should be tested by external patient datasets when possible.

Chapter 3

Computer modeling of patient AF for development of diagnostic tools

3.1 Introduction

As introduced in the previous chapter, non-invasive detection of AF mechanisms enable pre-operative screening and planning, making it the key to personalized treatment. Both ECGs and BSPMs acquire electric potential signals from multiple leads over the torso, and are important non-invasive mapping tools. The algorithms which measured the degree of organization in the atrial activity, had limited capacity in revealing underlying mechanisms, whereas for ECGi, while it was able to compute epicardial phase maps to identify AF driver [41], the accuracy for reconstructing epicardial activation patterns during AF episodes has not been validated [11].

In this chapter[1], we show how detailed AF models were built to assist the development of non-invasive diagnostic tools in two ways. For machine learning, first, we exploit the deductive nature of biophysical modeling, which provides mechanistic insights to show the reasoning for making predictions of ablation outcome using pre-ablation signals. Second, we use the models as digital twins of human atria to generate a large amount of synthetic data under various conditions for the development of diagnostic tools using machine learning. Specifically, we simulated AF based on several different underlying mechanisms.

Given that 1000 ms is a typical duration of f-waves during BSPM acquisition when using diltiazem to block the atrioventricular node and reduce ventricular interference, we designed computer models which incorporated several important AF mechanisms

[1]This chapter includes a part of the manuscript "Detection of focal source and arrhythmogenic substrate from body surface potentials to guide atrial fibrillation ablation" submitted to *PLOS Computational Biology*.

in a time scale of 1000 ms, as outlined in Fig. 3.1. We first considered if the AF was driven by a localized driver, defined as a fast and repetitive activity that causes fibrillatory conduction in the rest of the atria [154]. The determination of localized drivers is important, as it directly impacts the ablation strategies, whether to isolate local drivers or to extensively destroy the substrate. The localized driver is also commonly believed to be the main mechanism for early-stage AF patients [5]. Focal drivers contain stable focal sources (FS) with a dominant periodic rhythm [155], whereas reentrant sources, although spatially stationary, have higher heterogeneity in the cycle length (CL). For non-stationary drivers, multiple short-lived wavelets constantly evolve and collide in the atria. This spectrum, from the organized end to the disorganized end, shows the progressive nature of AF [156], which matches the AF electrophenotype spectrum in the recent review [9] that describes AF progression that shifts from FS-dominant to more substrate-susceptible. Because treatment of paroxysmal AF patients has been standardized, with PV isolation being the recommended treatment [5], we chose to model paroxysmal AF, and generated BSPMs and ECGs, driven primarily with FS and reentrant sources.

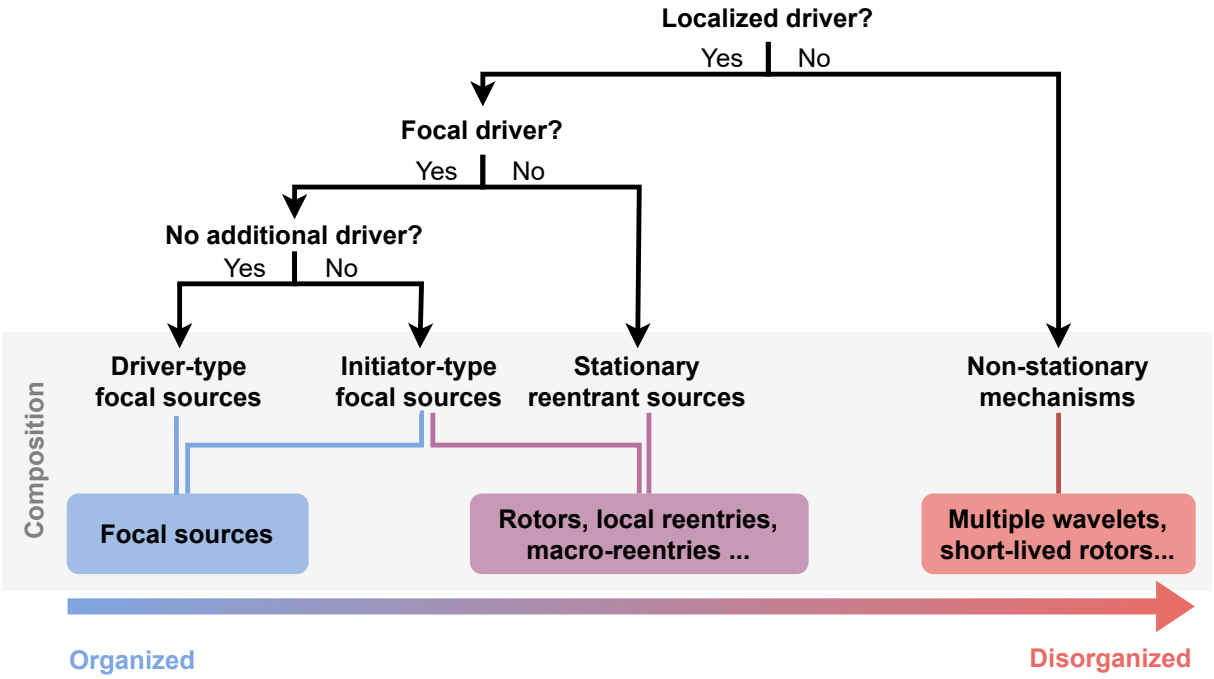


Fig. 3.1. Classification of AF mechanisms from organized to disorganized. The explanation of "Driver-type FS" and "Initiator-type FS" can be found in Section 3.2.3.

For development of non-invasive algorithms, we were especially interested in modeling FS, as they are an important trigger and initiator for AF. Because of this, electrical isolation or destruction of the FS, both PV and non-PV foci [157, 4], has been widely accepted as a standard AF treatment, especially for paroxysmal AF patients [5]. However, an arrhythmogenic substrate, which is not usually targeted by focal ablation procedures, can still render the atria susceptible to AF. Technologies that enable non-invasive identification of FS and the substrate can be valuable for personalized treatment.

Another reason to model FS was that, although high counts of atrial ectopic beats measured by 24-hour or longer Holter ECGs in asymptomatic subjects have been associated with a higher risk of AF incidents [160, 161, 162], ectopic beats are also markers of other cardiac conditions associated with AF, such as stroke, the causality from frequent ectopic beats to AF incidents could not be established [163], indicating a need for other markers which enable mechanism-based prediction. The identification of malignant ectopic beats with likely AF progression potentially informs early intervention, which admits higher treatment success rates [164], as chronic AF results in tissue remodeling and fibrosis [37] which stabilize AF.

3.2 Bilayer atrial model

We represented the patient atrial mesh using a bilayer model [106], which captures the different fiber orientation and electrical properties of the epicardial and endocardial surfaces of the atria, with an intermediate computational complexity between that of a surface mesh and that of a volumetric mesh.

3.2.1 Geometry

Patient bi-atrial meshes were made from late-gadolinium enhanced magnetic resonance imaging scans of five AF patients from the CHU Bordeaux [165, 107]. The surface of the intra-atrial blood pools was used as the endocardial surface and the epicardium was projected 100 μm in the normal direction. Meshes were composed of on average 489k nodes, and 1.18M elements, primarily triangular elements between nodes, and line elements between endocardial and epicardial nodes, such as between the pectinate muscle and RA epicardium. The average edge length was 318 μm (Table 3.1). Inter-atrial structures were added to the meshes on two atrial chambers, such as between fossa ovalis nodes on the LA and on the RA using a standardized atrial coordinate. In the LA, the endocardial and epicardial surfaces were made 100 μm apart in the surface normal direction, and the RA endocardium only includes pulmonary muscle and crista terminalis [106]. Their orientations were aligned, and their centers of mass had standard deviations of 1.3, 1.8 and 1.0 cm in the x, y, z directions, respectively.

The atrial regions are shown in Fig. 3.2. The fiber orientation (in Fig. 3.3) was transferred from Labarthe et al. [106] based on findings from histological study by Ho et al. [166, 167] to the five meshes using UAC [165, 107].

To obtain the locations of BSPM leads, the meshes made from computerized tomography imaging of the atria and 252-lead vest (CardioInsight, Medtronic, MN) from a sixth patient were also exported. The atrial mesh of the first patient was registered

Table 3.1. Statistics of patient atrial meshes for simulations.

	Number of elements	Number of nodes	Edge length (μm)
Mesh 1	1191111	492799	320
Mesh 2	1445899	595997	318
Mesh 3	963281	402343	318
Mesh 4	1270575	528589	319
Mesh 5	1040902	427664	319

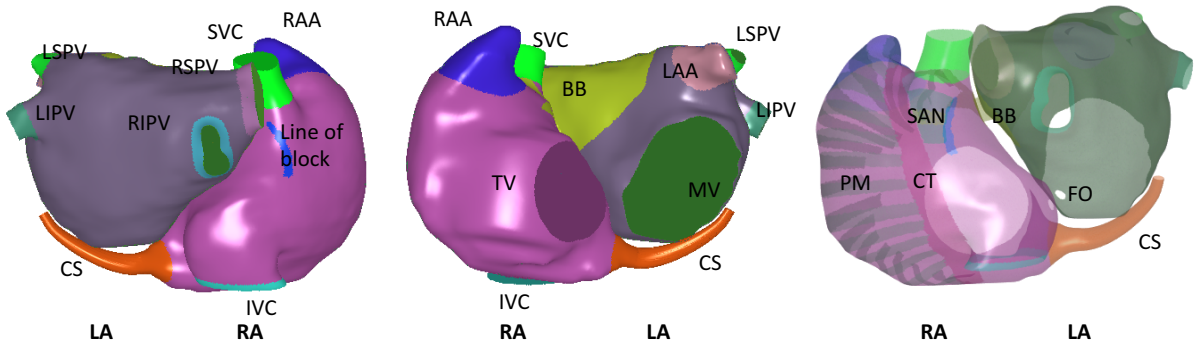


Fig. 3.2. Atrial regions in the model on posterior (left) and anterior (middle) views, with inter-atrial structures (right) containing CS, BB, FO. SAN: Sino-atrial node, LA: left atrium. RA: Right atrium. CS: coronary sinus. BB: Bachmann's bundle. FO: fossa ovalis, Right Superior Pulmonary Vein: right superior pulmonary vein. RIPV: right inferior pulmonary vein. Left Superior Pulmonary Vein: left superior pulmonary vein. LIPV: left inferior pulmonary vein. PM: pectinate muscle. CT: cristi terminal. LAA: left atrial appendage. RAA: right atrial appendage. SVC: superior vena carva. IVC: inferior vena carva. TV: tricuspid valve. MV: mitral valve.

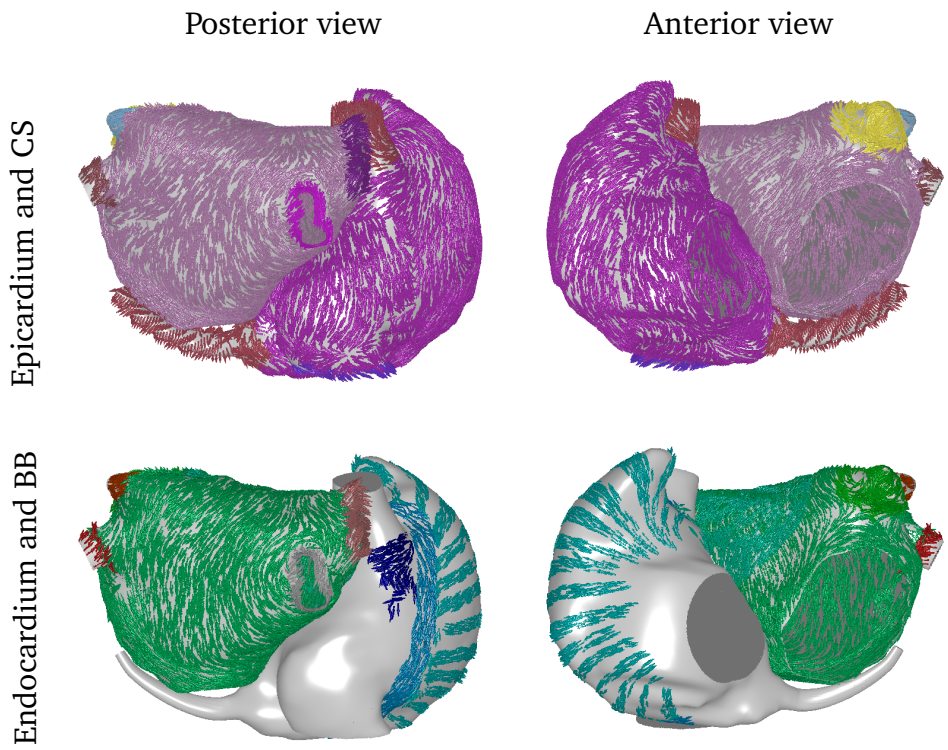


Fig. 3.3. The fiber orientation of the atria, with regions shown in different colors. CS: coronary sinus. BB: Bachmann's bundle.

with that of the sixth patient by matching the centers of mass of the two atria by linear translation, i.e. rotation by 90 degrees and translation of the three Euclidean axes. All five virtual patients then shared the same vest leads in Fig 3.4.

Universal Atrial Coordinates (UAC)

To refer to locations in different patient meshes in an automatic and uniform way, all nodes in the mesh were assigned two coordinates $(\alpha_{side}, \beta_{side})$, where $side \in \{LA, RA\}$, $\alpha_{side} \in [0, 1]$, and $\beta_{side} \in [0, 1]$, from UAC [109], with the effects in Fig. 3.5. The UAC system assigns coordinates of each node according to the relative locations to important anatomical structures, by computing Laplace-Dirichlet solution between defined boundaries in the atria. It is done by setting one boundary as 0 as set another boundary as 1, so that a linear transition of the values between two boundaries can be obtained. The boundaries include important atrial structures such as superior PVs, fossa ovalis and mitral valve on the LA, and tricuspid valve, superior vena cava, inferior vena cava and fossa ovalis on the RA, and geodesic lines between these structures.

The Laplace-Dirichlet solution is also needed to separate anterior from posterior areas on the LA, and between septal and lateral areas for RA. This can be done by first separating the two areas with an enclosed boundary line, and setting all nodes

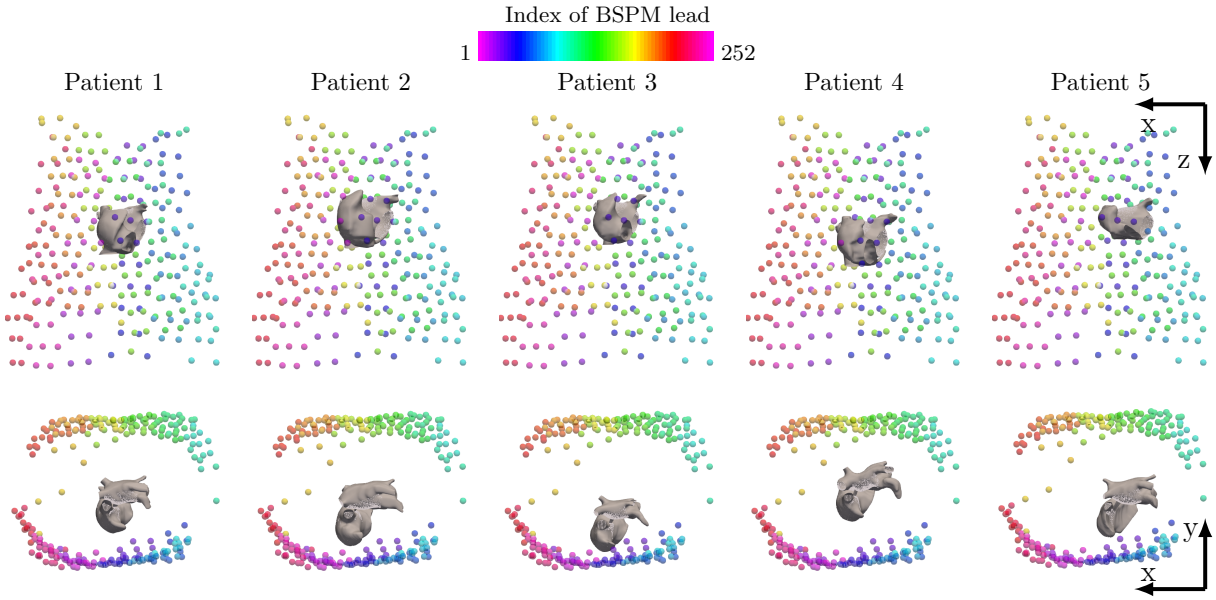


Fig. 3.4. Meshes of five patient atria, and locations of 252 vest leads to compute BSPMs. Each column shows the data for one patient. The top row shows the front view and the bottom row shows the top view. Inter-patient variations of atrial shapes and positions are visible.

on this boundary line as 0, and setting another point inside the boundary to 1. From the Laplace-Dirichlet solution, the nodes with positive values are inside the boundary, whereas the nodes with value 0 are outside the boundary.

3.2.2 Electrophysiology

Electrical activity was described in the bilayer model by the monodomain equation. The monodomain formulation describes the cardiac tissue as a single conduction domain, by assuming the anisotropy ratios in the intra- and extra-cellular conductivities are related by a scalar λ by $\sigma_e = \lambda\sigma_i$. The monodomain equation has been found to well approximate electrograms computed by the full bidomain equations [168]. It is described by the following reaction-diffusion equation [169]:

$$\nabla \cdot (\sigma_m \nabla V_m) = \beta C_m \frac{\partial V_m}{\partial t} + \beta I_{ion} - \beta I_{stim} \tag{3.1}$$

On the left hand side of Eq. (3.1), there are the transmembrane voltage V_m (unit: mV) and the equivalent monodomain conductivity σ_m (unit: $mS/\mu m$), which is the harmonic mean of the extracellular and intracellular conductivity tensors:

$$\sigma_m = \frac{\lambda}{1+\lambda}\sigma_i \tag{3.2}$$

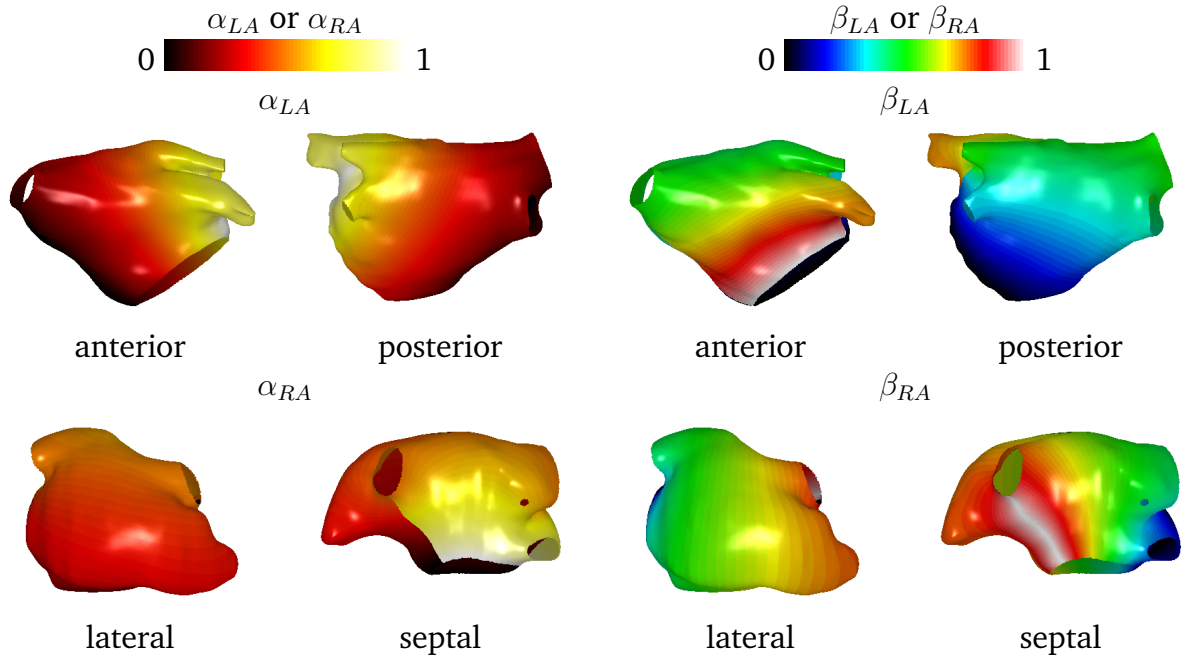


Fig. 3.5. Example UAC system coordinates (including α_{LA}, β_{LA}, α_{RA} and β_{RA}) on the mesh of the Patient 5, computed based on the positions of important atrial structures such as PVs, tricuspid valve and mitral valve. The computation result was obtained from Roney et al. [109].

The intracellular (or the extracellular) conductivity per unit length σ_i (or σ_e) (unit: $mS/\mu m$) measures the ability of the intracellular (or the extracellular) domain to conduct electric currents. The tensors are anisotropic with distinct longitudinal (g_{il}), and transverse conductivities (g_{it}).

The right hand side of Eq. (3.1) contains three currents [22]. The first type of current refers to the capacitive current, I_{cm}, due to the membrane acting as a capacitance that separates the opposing charges inside and outside the cell. Given the membrane capacitance per unit surface, C_m (unit: $\mu F/cm^2$), it follows

$$I_{cm} = C_m \frac{\partial V}{\partial t}$$

The second type is the ionic current I_{ion} (unit: $\mu A/cm^2$) which is the net membrane current produced by the ionic model chosen to represent cellular behaviour. We used the Bayer et al. [170] modifications of the CRN model [27] with included addition of ACh currents, where the ionic current has the following form:

$$\begin{aligned} I_{ion} = {} & I_{Na} + I_{K1} + I_{to} + I_{Kur} + I_{Kr} \\ & + I_{Ks} + I_{Ca,L} + I_{p,Ca} + I_{NaK} + I_{NaCa} \\ & + I_{b,Na} + I_{b,Ca} + I_{K(ACh)} \end{aligned} \tag{3.3}$$

where each term is a non-linear function of V_m resulting from an ion species traversing

the corresponding ion channel. In addition, Eq. (3.1) takes an external stimulus current density I_{stim} (unit: $\mu A/cm^2$) into consideration, with the negative sign signifying an inward current. All of the current densities need to be multiplied by the surface-to-volume ratio β (unit: μm^{-1}) of the cell to match the unit of the left hand side.

Forward computation of BSPM and ECG

Extracellular potentials were computed using a classical ϕ_e-recovery method [22] at the field potentials of the vest leads, which constitutes the most straight-forward way to compute far-field extracellular potentials Φ_e at the BSPM lead,

$$\Phi_e = \frac{1}{4\pi\sigma_b} \int_\Omega \frac{I_m}{r}\, d\Omega \tag{3.4}$$

where σ_b is the bath conductivity, the scalar r denotes the distance between the source and the electrode position, and I_m is the current density derived from the transmembrane potential V_m obtained from the monodomain solution, using $I_m = \nabla\cdot(\sigma_\mathbf{i}\nabla V_m)$ [171], integrated over the full tissue domain Ω.

To recreate clinical recordings, potentials were referenced to the Wilson Central Terminal. 12-lead ECG potentials were calculated from a subset of the BSPM potentials. Specifically, the potentials of the three limb leads (V_L, V_R and V_F) and the six precordial leads (V_1, V_2, ..., and V_6) were taken directly from the vest leads that best fit the locations of the corresponding ECG electrodes. The resultant 12-lead ECGs consisted of the potentials of the six precordial leads, and six augmented leads (I, II, III, aV_R, aV_L and aV_F), which were calculated using Eqs. (2.3)–(2.8).

The heterogeneity in conductivity between the heart and the torso mostly impacts the amplitude and, to a much smaller extent, the morphology of body surface potentials calculated by forward computation [172]. In an ECGi study, the pattern of the reconstructed epicardial potentials obtained from ECGi using a homogeneous torso was comparable to using a heterogeneous torso [173]. Both studies suggest that a homogeneous torso can be used to generate body surface potentials, as long as errors of potential amplitude are considered during interpretation. Therefore, for computational efficiency, heterogeneous conductivities of different organs were not included in the forward computation of BSPMs and ECGs, so we did not have to construct a torso finite element model.

Simulator

The monodomain solution and the forward computation were solved by the finite-element method [106] using the CARP software [169], with a time resolution of 20

μs. The CARP software runs in parallel on top of PETSc [174, 175, 176] with Open MPI software [177] (https://www.open-mpi.org/). We tested the run time of a 1000 ms long simulation of about 500,000 vertices. On an Intel® Xeon® Gold 6140 processor, it took on average 70 min to finish the simulation when running with 16 processes. In the supercomputer, when the program was assigned 448 cores, it took on average 3.5 min to finish the simulation.

Electrical heterogeneity

Ionic models were adopted from previous work by Roney et al. [165, 107], which followed Bayer et al. [117] for ionic models of non-PV structures. Essentially, in the LA, the maximal conductance of the transient outward K^+ current g_{to}, the ultrarapid delayed rectifier K^+ current g_{Kur} and the L-type inward Ca^{2+} current $g_{Ca,L}$ were decreased by 50%, 50% and 70%, respectively from the Courtemache-Ramirez-Nattel (CRN) model [27], which matched the APD morphology, APD restitution curves as well as the activation latency in AF patients, according to Krummen et al. [178]. The maximal conductance of the fast inward Na^+ current g_{Na} from the CRN model was doubled by Labarthe et al. [106] to be kept within a physiologically realistic range on the cytoplasm for the atrial tissue. The maximal conductance of the inward rectifier K^+ current g_{K1} was downgraded by 20% to obtain a physiologically realistic baseline APD as Krummen et al. [178]. For other structures, regional heterogeneity was added by references to Seemann et al. [179] and Aslanidi et al. [180]. For the appendage structures, the maximal conductance for the rapid delayed rectifier K^+ current g_{Kr} was also adjusted to 1.6 times higher in the LA appendage compared to the RA appendage as Bayer et al. [117] which preserved the relative values of ionic current conductivities between LA, LA appendage and RA appendage regions in Dorn et al. [181]. PV parameters were added from CRN based on the experimental data of Krueger et al. [182]. The electrical properties of the sino-atrial node were added according to Fedorov et al. [183]. Finally, in order to enable ACh regulation, for the LA and its appendage, a human non-fibrotic ACh-regulated AF ionic model by Bayer et al. [170] was used, which adjusted the ACh-activated potassium current $I_{K(ACh)}$ in the Kneller's model [184] to match the electrophysiology of LA and LA appendage in human, while leaving the other parameters in the CRN model unchanged.

The regional conductivities in the bilayer model (Table 3.3) were obtained from Labarthe et al. [106], which minimized the error with respect to the activation times reported in Lemery et al. [185] Table 3. Because the ionic models from the previous work had the same conduction velocity with the non-AF atrial cells, in order to match the conduction velocity slowing of the LA and LA appendage for AF patients for the ACh-regulated AF ionic model [170], we scaled the conduction velocity of the crista termina, pectinate muscle, Bachmann's bundle and coronary sinus and the RA myocardium at both longitudinal and transverse directions by a factor of 0.336, which is the geometric

Table 3.2. The maximal conductance (unit: pS/pF) for the modified CRN model in each region. SAN: Sino-atrial node, LA: left atrium. RA: Right atrium. CS: coronary sinus. BB: Bachmann's bundle. PV: pulmonary veins. PM: pectinate muscle. CT: cristi terminal. LAA: left atrial appendage. RAA: right atrial appendage.

Anatomical region	g_{Na}	g_{to}	g_{Kr}	$g_{Ca,L}$	g_{K1}	g_{Kur}	Source
Default	15600	82.6	29.4	37.1	72.0	$\times 0.5$	[178, 106]
LA	15600	82.6	47.1	37.1	72.0	$\times 0.5$	[179, 180, 106, 170]
sino-atrial node	15600	165.2	29.4	123.8	72.0	$\times 0.9$	[179, 180, 106]
BB, CT	15600	105.8	29.4	62.0	72.0	$\times 0.5$	[179, 180, 106]
LAA	15600	56.2	47.1	39.3	72.0	$\times 0.5$	[181, 106, 117, 170]
RAA	15600	56.2	29.4	39.3	72.0	$\times 0.5$	[181, 106, 117]
PVs	15600	102.3	20.6	27.8	60.3	$\times 0.5$	[182, 106]
Original CRN	7800	165.2	29.4	123.8	90.0	$\times 1$	[27]

mean of conductivity scaling in the longitudinal and transmural directions when scaling from non-fibrotic non-AF to non-fibrotic AF models in Bayer et al. [170]. The conduction velocities of four unidirectional connections with the RA tissue were assigned a very high conduction velocity to mimic the activity of the sino-atrial node suggested by Fedorov et al. [183]. For the line of block and inferior vena cava which remained inactivated during a sinus rhythm, we assigned a small conductivity of 0.001 S/m.

Table 3.3. The longitudinal and transverse conductivities (g_l and g_t, unit: S/m) in each region. SAN: sino-atrial node, LA: left atrium. RA: Right atrium. CS: coronary sinus. BB: Bachmann's bundle. PM: pectinate muscle. IVC: inferior vena carva. PV: pulmonary veins. CT: cristi terminal. LAA: left atrial appendage. RAA: right atrial appendage.

Type of elements	Anatomical region	g_l	g_t
Triangular elements	All myocardium	0.107	0.045
	CT	0.165	0.025
	BB	0.236	0.061
	PM	0.234	0.010
	sino-atrial node	0.392	
	CS	0.134	
	Line of block, IVC	0.001	
Linear bilayer connections	CT-RA	1.054	
	LA endo-epi	0.326	
	RA-BB-RA	3.000	
	PM-RA	1.054	
	SAN-RA	3000	

3.2.3 Simulations of AF mechanisms

To generate synthetic AF episodes resembling those in paroxysmal AF patients, we modeled AF driven by FSs and reentrant drivers, and applied focal ablation to further classify the FS episodes according to the AF sustainability after the ablation. We marked those with sustained AF after focal ablation as *Initiator-type FS (IFS)*, and *Driver-type FS (DFS)* otherwise. IFS induced AF episodes were marked as ***FS-induced AF (FAF)***, and AF episodes initiated by localized reentrant sources were marked as ***Reentrant AF (RAF)***. In Fig. 3.1, DFS corresponds to the organized end, and IFS corresponds to the mid-point of the spectrum. RAF and FAF include both stationary reentrant sources at the mid point, and non-stationary mechanisms at the disorganized end. The composition of simulations is shown in Fig. 3.6.

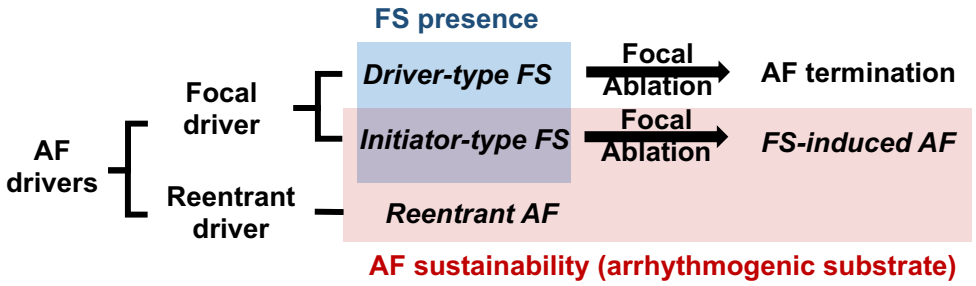


Fig. 3.6. The composition of simulations, included Driver-type FS, Initiator-type FS, FS-induced AF and Reentrant AF, with the classification of "FS presence" (shaded in blue), and "AF sustainability" (shaded in red). We modeled FS and reentrant sources (Reentrant AF), and removed the FS to test if its removal terminated AF, in order to decide if it was a Driver-type FS or an Initiator-type FS. "FS presence" includes all FS driven episodes, whereas "AF sustainability", hinting at the presence of an arrhythmogenic substrate, included episodes where AF did not terminate by focal ablation. Note that the Initiator-type FS is in both boxes of "FS presence" and "AF sustainability".

Focal drivers

To produce a steady electrical state before the FS were induced, the atria was first left resting for 1000 ms, and the sino-atrial node was paced five times with an CL of 700 ms, where the strength of the stimulus current was 44 $mA/\mu A^3$ for 2 ms. Starting at 3500 ms, single-site repetitive pacing at specified locations was performed with CLs of $120, 150, \ldots, 270$ ms, the clinically observed range of FS CLs [4], over a period of 3000 ms to reach a stable state for the FS. The strength of the pacing current stimulus was 200 $mA/\mu A^3$. FS sites were located on the myocardial sleeves of the four PVs as well as at 32 locations of the LA and the RA at combinations of (α, β) with $\alpha_X, \beta_X \in \{0.2, 0.4, 0.6, 0.8\}$ and $X \in \{RA, LA\}$. Each FS was identified by its focal site, CL, as well as [ACh].

Focal ablation to test additional drivers in the presence of a focal source

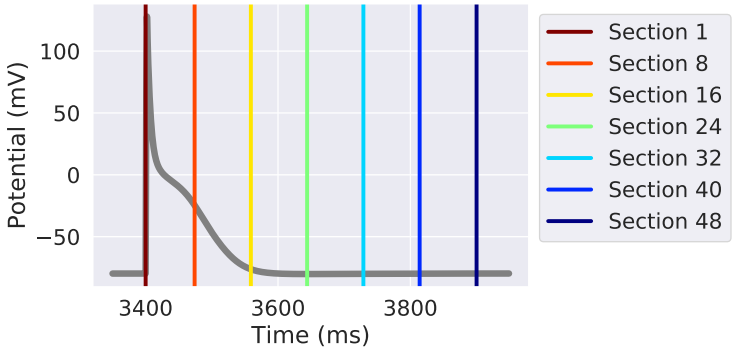
We tested whether each FS induced self-sustained AF, defined as lasting for 1000 ms after stopping the FS, like in focal ablation which destroys or isolates the FS. If AF stopped after focal ablation, the FS was the driver and denoted as a DFS. Otherwise, the FS induced AF, and was denoted as an IFS. The 1000 ms AF segment after the pacing stopped for an Initiator-type was labelled as a FAF. The last 1000 ms segments of constant FS pacing, as well as the 1000 ms segments of FAF after pacing stopped, were added to our synthetic dataset.

Initiating Reentrant drivers

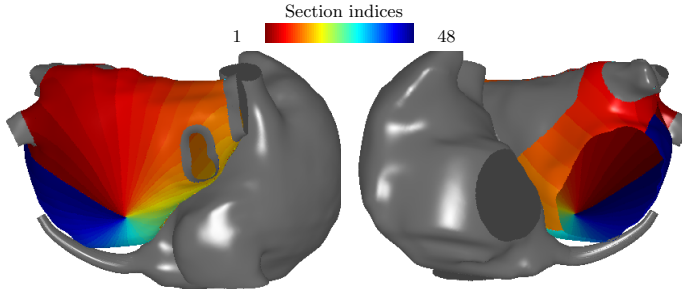
In clinical and experimental settings, fast or decremental pacing at one location can be used to induce a reentrant driver [38]. However, there is no control on where the reentrant circuits form. Therefore, we used a phase distribution method [186] to locate the rotor at specified locations. In our experience, it is easier to control the spatial location of the rotors using the phase distribution than the cross-field stimulation method, as the cross-field stimulation method requires a carefully timed delay between S1 stimulus and S2 stimuli to have a unidirectional conduction block and the rotor core at intended location [187]. In a three-dimensional atrial model with spatial heterogeneity in electrical properties, a time-consuming trial-and-error method must be used to find the exact S2 delivery.

To use the phase distribution method for initiating reentrant sources, we added a non-excitable core region by setting the resting potential to -81 mV within the region, and effectively zero conductivity, with the radius of the region being 0 cm, 0.5 cm or 1 cm and located at the same 32 sites as the focal sites. The atrial surface was divided into 48 equal sectors around the core to determine the initial conditions for each sector. The states for the sectors were taken from 48 time instances equally sampled from a single-cell action potential during a full depolarization cycle, with a 2 ms transmembrane current stimulus of 80 $\mu A/cm^2$, matching the parameters of the atrial site. An illustration of the process is shown in Fig. 3.7.

Only those simulations with sustained activity for more than 3000 ms were kept for further investigation, and the last 1000 ms segment of the resultant AF was marked as an RAF, and added to our synthetic dataset. This interval of 3000 ms was adopted so that their dynamics were regular, and comparable to FS. Additionally, a rotor lasting for more than 3000 ms was also likely to be sustained for a long time, say, 80 seconds [188].



(a)



(b)

Fig. 3.7. An example setup to initiate a reentrant source around an LA site $(\alpha_{LA} = 0.2, \beta_{LA} = 0.2)$ **on the mesh of Patient 2.** The areas initialized to rest are marked in gray. To initiate a rotor on the LA, the LA body was split into 48 sectors, marked by 48 different colors in (b), around a non-excitable core. In this case, the non-excitable core had a radius of 0. The initial states were generated from an action potential, with selected states shown in (a).

Acetylcholine (ACh) distribution

To add APD heterogeneity to the model, we also simulated islands of ACh release ([ACh] $= 0.1\mu$M) on the LA surface and the LA appendage [170], where 60, 80 and 100 2mm-radius ACh islands were randomly distributed at areas of α_{LA} of $0-0.33, 0.33-0.67$ and $0.67-1$ and β_{LA} of $0-1$, respectively. The ionic model and the conduction velocity for the ACh islands are the same as those of LA, but an ACh sensitive potassium channel was added to the model. An example ACh distribution is shown in Fig. 3.8. Both FS and RAF sources were applied in conditions with ([ACh]$=0.1\mu$M) and without ACh.

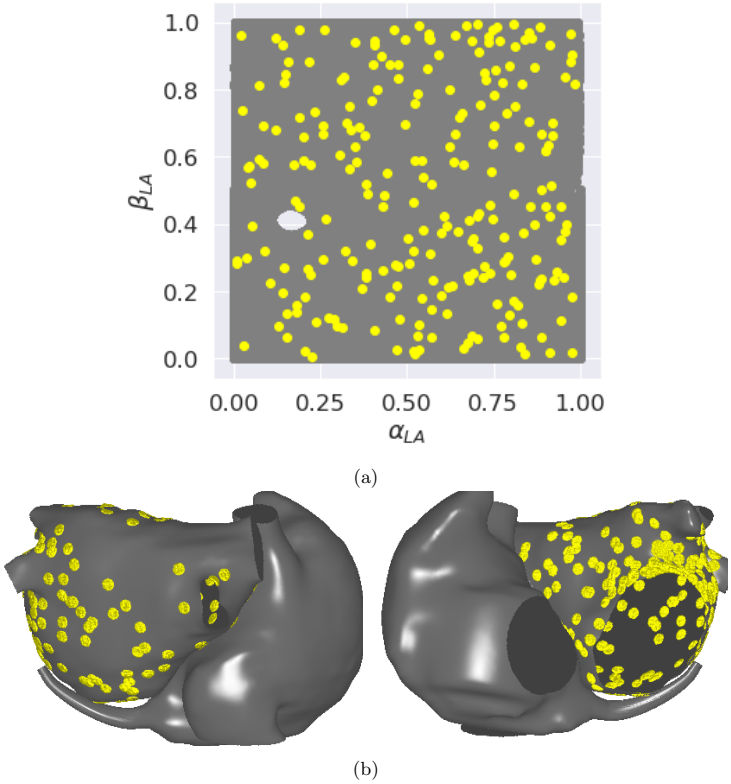


Fig. 3.8. An example ACh distribution with (a) two-dimensional mapping, and (b) three-dimensional representation on the atrial mesh of Patient 2. The other areas without ACh islands are marked in gray. As mentioned in the text, fixed numbers of 60, 80, and 100 2mm-radius ACh islands (in yellow) were randomly distributed at areas of α_{LA} of $0-0.33, 0.33-0.67$ and $0.67-1$ and β_{LA} of $0-1$, respectively. Note that in (a), only the positions but not the sizes of ACh islands were indicated.

3.3 Data analysis

In total, there were 5 (patients) $\times$ 36 (LA, RA and PV foci) $\times$ 6 (CLs) $\times$ 2 ([ACh]) $=$ 2160 FS cases, and a small number (60) of them were not launched as their UAC coordinates of foci referenced locations outside of the atrial tissue. This left 2100 FS

simulations, with 900 LA, 960 RA and 240 PV FSs. The odds of inducing self-sustained AF for at least 1000 ms from FS under a specific condition, measured as the proportion of IFS, was termed *AF inducibility*. The AF inducibility test identified 1772 DFSs, and 328 IFSs, which entailed another 1772 discarded simulations and 328 FAF simulations. We initiated 5 (patients) $\times$ 32 (LA and RA core sites) $\times$ 3 (core radii) $\times$ 2 ([ACh]) = 960 RAF simulations, with 549 sustaining for more than 3000 ms being preserved. These 1772 (DFS) + 328 (IFS) + 328 (FAF) + 549 (RAF) = 2977 segments were used as study data in the following steps.

3.3.1 Calculation of rotor duration

In order to track rotors in these simulations, we calculated phase singularity (PS) points, i.e., the cores of the rotors. We first filtered the transmembrane voltage with a fourth-order 15 Hz low-pass Butterworth filter, removed the DC component with a 0 Hz notch filter, computed phase by applying the Hilbert transform, and then implemented the Iyer-Gray method [189] which identifies PS points as those around which a contour integral of the phase equals $\pm 2\pi$. This contour was defined by the three nodes of a triangular element, and all the elements can be processed simultaneously. To calculate all rotors and sustained rotors, we kept rotors lasting for at least 10 ms and at least 100 ms, respectively, assuming the PS point moves less than 2 mm within 1 ms. The total duration with at least one such rotor was also computed for analysis.

3.3.2 Calculation of local activation and repolarization time

Local activation and repolarization times were marked by the upward-sloped intersection of -30 mV and the downward-sloped intersection of -70 mV of the transmembrane voltage since the time of a focal discharge, respectively. The fixed cut-off value helped to process the transmembrane voltage in a vector form. We then calculated the APD by subtracting the repolarization time from the local activation time.

3.3.3 Approximation of effective refractory period (ERP)

ERP was estimated as the maximal FS CL whereby the FS encountered a complete conduction block at tissue surrounding the focus. Searched on the set of prescribed FS CL*s*, ERP was estimated between the maximal CL and the minimal CL where a complete conduction block did or did not occur near the pacing site.

3.4 Arrhythmia induced

3.4.1 Simulation effects

Although we initialized simulations with one FS or one reentrant source, our induced AF episodes may include one or many simultaneous sustained ($>$ 100 ms) rotors, as shown in Fig. 3.9.

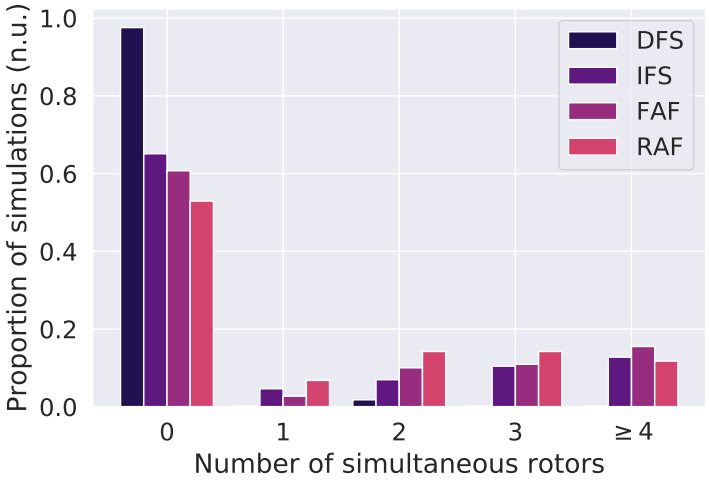


Fig. 3.9. A histogram of the number of simultaneous sustained ($>$ 100 ms) atrial rotors for simulations of DFS, IFS, FAF and RAF.

For visualization, selected frames of an IFS and a DFS are shown in Fig 3.10, where a sustained rotor lasting for 547 ms in total can be seen during the IFS episode in (a), but there was no sustained rotor in the DFS episode in (b). The effects of focal ablation on these FS are shown in Fig. 3.11, where the activation continued on the IFS as a form of macro-reentry between the LA and the RA via the coronary sinus ostium in (a), but not on the DFS. The RAF, including a rotor in (a) and a reentry around an inexcitable core of radius 1 cm in (b), are shown in Fig. 3.12. The results of adding ACh regulation are shown in Fig. 3.13 where small wavelets (shown in red arrows) were located in areas near LA appendage with ACh islands densely distributed.

For forward modeling, the examples of ECGs of an IFS, a DFS and an RAF is shown in Fig. 3.14, where the amplitude was normalized to show the organization. Visually, the DFS had the most organized morphology, followed by the IFS, and the RAF was the most disorganized.

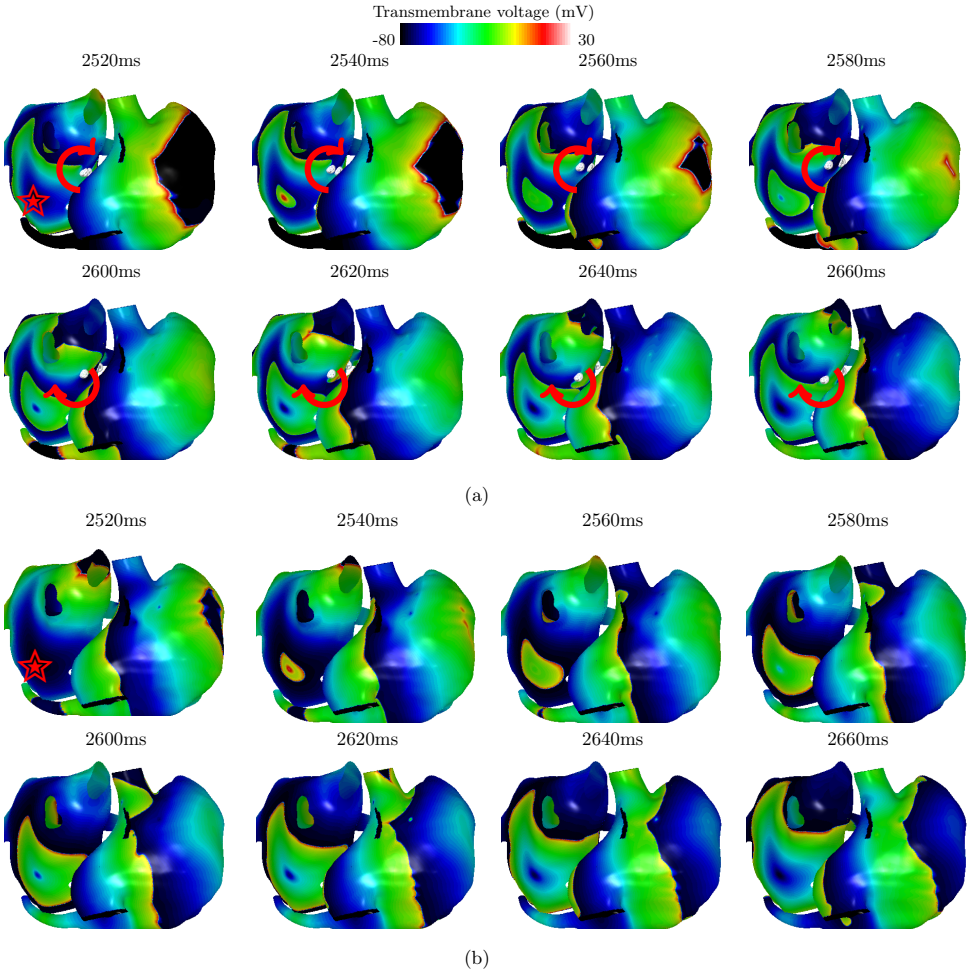


Fig. 3.10. Simulations on the atrial mesh of Patient 2 of (a) an IFS with a CL of 180 ms at a LA focal site ($\alpha_{LA} = 0.2, \beta_{LA} = 0.2$, red star), and (b) a DFS with a CL of 210 ms at the same site (red star). The white spheres in (a) show PS points, and the red arrows show the wavefront movement of rotors. The time since the first firing of the FS was indicated above each frame. A sustained rotor of a total duration of 547 ms was presented in (a), but no rotor was detected in (b).

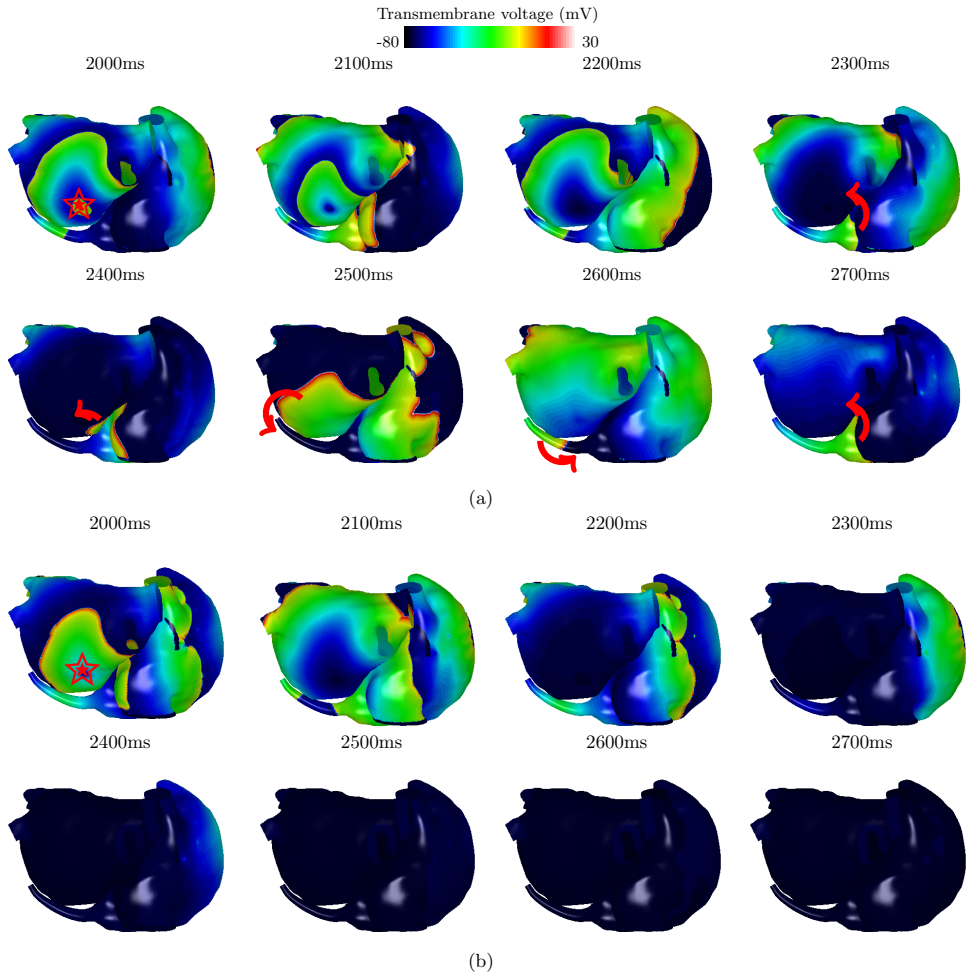


Fig. 3.11. Effects of removing the FS for (a) the IFS with a CL of 180 ms in Fig 3(a), where the AF continued after pacing stopped, and for (b), the DFS with a CL of 210 ms in Fig 3(b), where AF died out from 2400 ms, on the atrial mesh of Patient 2. The red stars mark the locations of the FS activated right before 2000 ms (at 1980 ms for (a) and at 1890 ms for (b)). The red arrows showed the movement of rotor wavefronts. The time was counted from the first pacing of the FS. A macro-reentry going through the coronary sinus can be seen from 2300 – 2700 ms on (a).

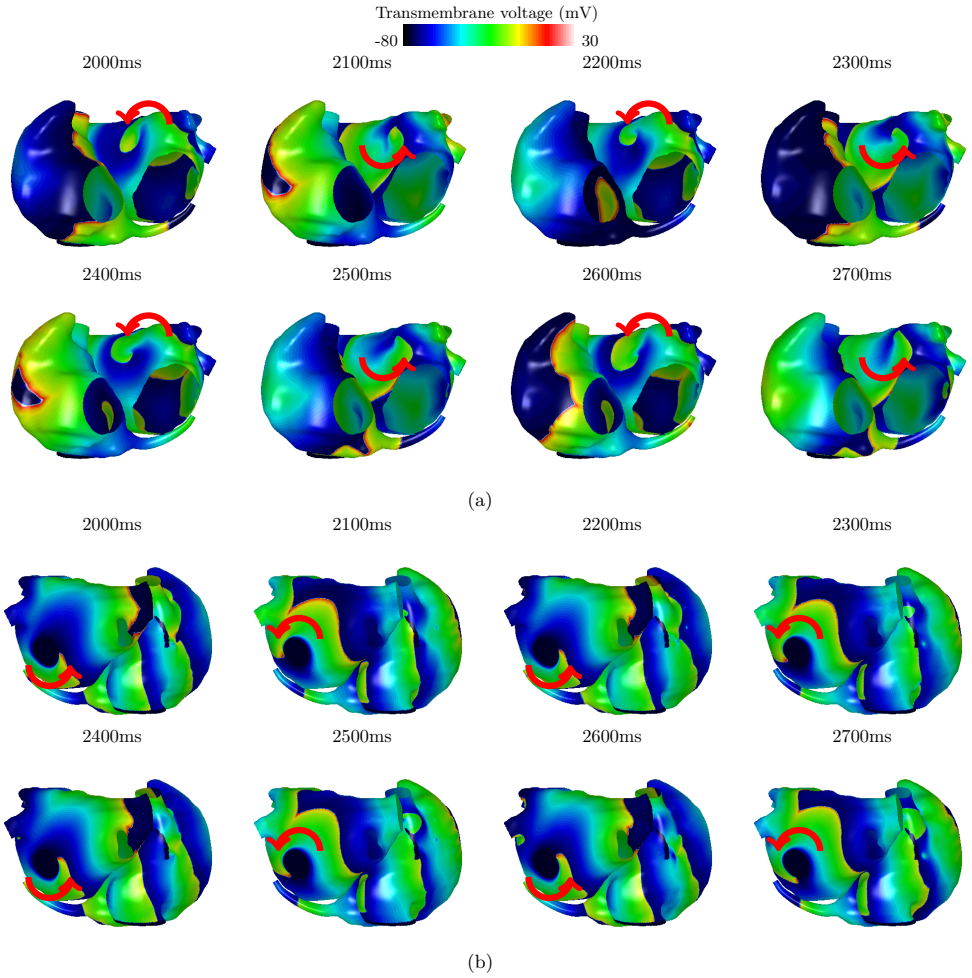


Fig. 3.12. Simulations of RAF, including (a) rotor, and (b) reentry with a conduction block of radius 1 cm on the atrial mesh of Patient 2. The red arrows indicated the directions of wavefront propagation, and the rotational sources can be clearly seen. The time was counted from the initiation of the AF.

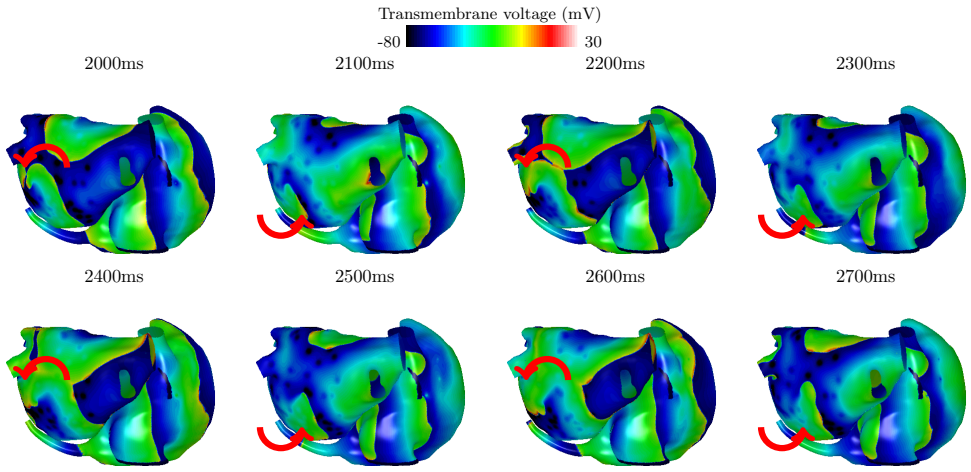


Fig. 3.13. Simulation of the effect of ACh, on the atrial mesh of Patient 2. The red arrows indicated the directions of wavefront propagation. The red arrows indicate the directions of wavefront propagation. The time was counted from the initiation of AF. A local rotor appeared close to the base of the posterior left atrial wall where ACh islands were distributed. The potentials were also more heterogeneous with the introduction of ACh.

3.4.2 Distributions of AF inducibility from FS

We systematically changed the CL and the location of FS, and the concentration of ACh ([ACh]), across five patient atria, and calculated the AF inducibility and the sustained rotor and all rotor duration for each case of CL, focal site and [ACh] on each patient. The means and 95% confidence intervals across five patients are shown in Fig. 3.15, and the increase of AF inducibility from FS by adding ACh regulation is shown in Table 3.4.

AF inducibility was the highest from the PV FS, followed by the LA FS. The highest AF inducibility was observed from FS with a CL of 180 ms from the PVs and from the LA, with means 0.95 and 0.88, respectively, and with a CL of 210 ms from the right atrium (RA), with mean 0.24. The CL range with non-zero AF inducibility was also larger for FSs from the PVs and the LA than in the RA. ACh regulation increased AF inducibility from FSs with CLs of 150 – 210 ms on the LA, and from FS with a CL of 150 ms on the PVs, but not for FS from the RA. Over five patient meshes, FSs at a CL of 210 ms showed a low average AF inducibility of 26.0% with a moderately high inter-patient variability (standard deviation $\pm 18.5\%$), likely due to the differences in the geometry between five atria.

The average AF inducibility for each FS over different atrial focal sites across five virtual patients is shown in Fig. 3.16. At the LA appendage and left PVs, where the ACh is densely distributed, the AF inducibility increased with ACh regulation in the CLs of 150 to 210 ms.

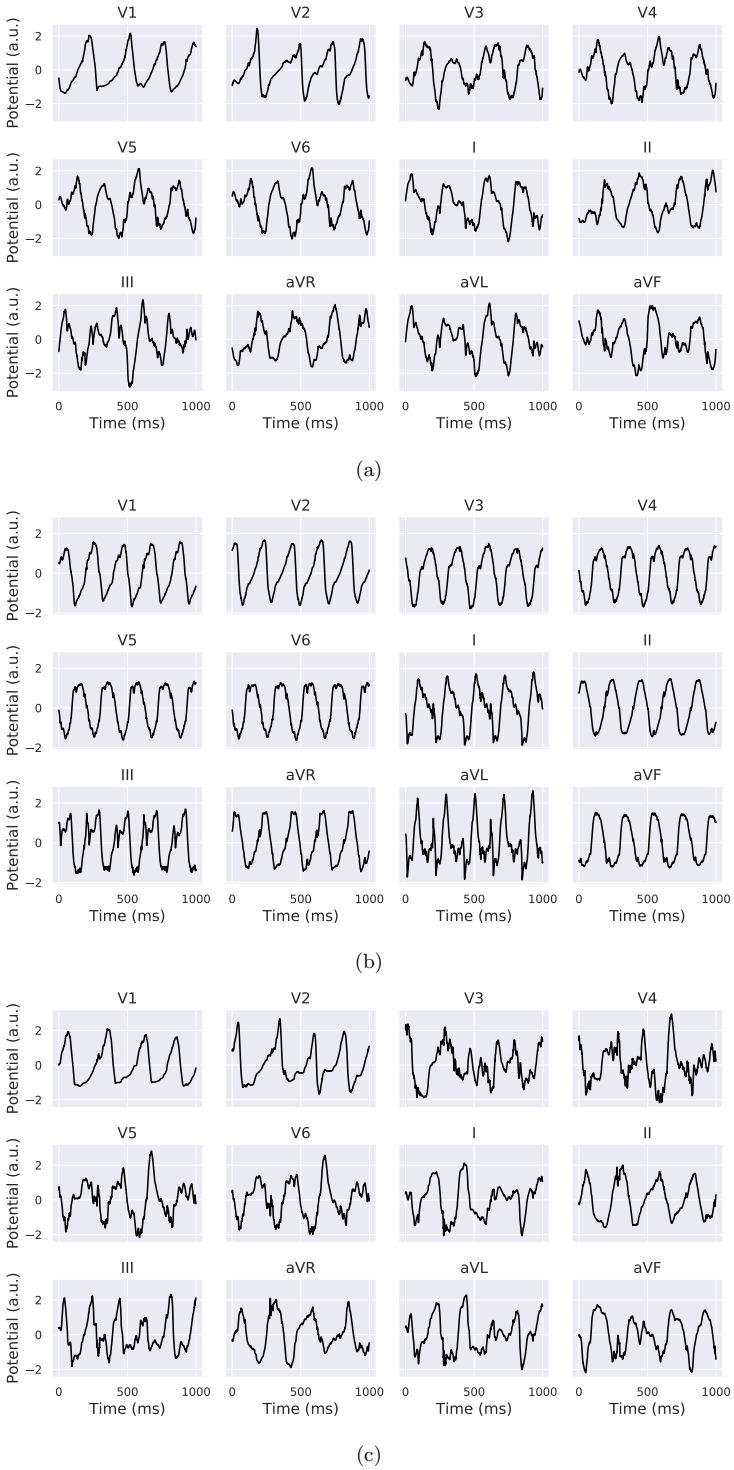
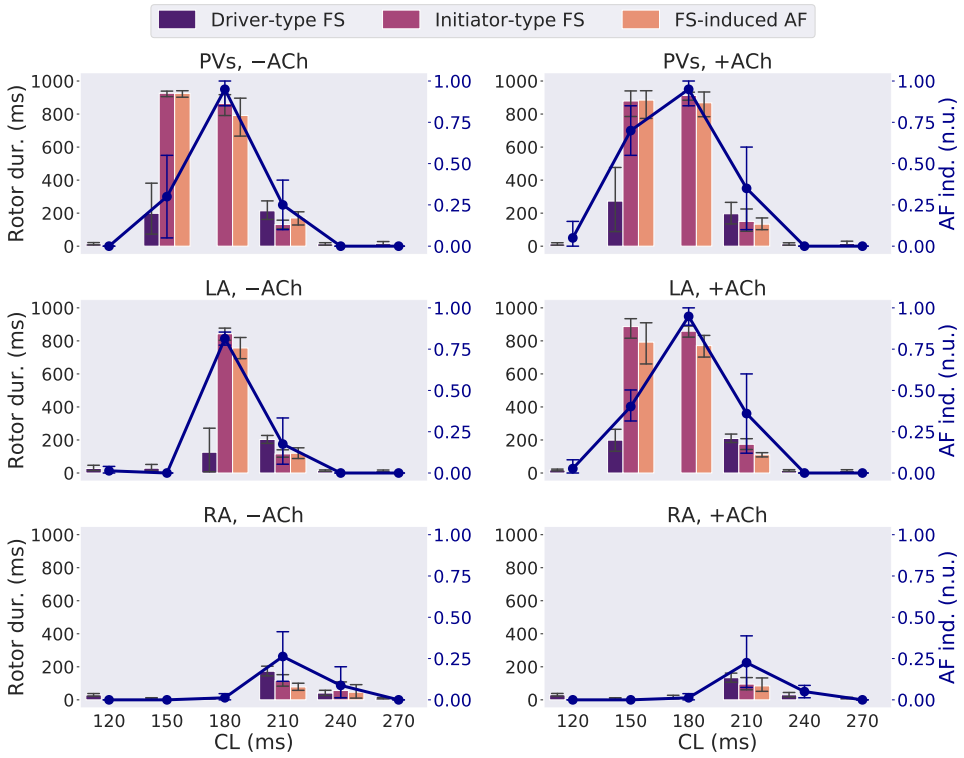
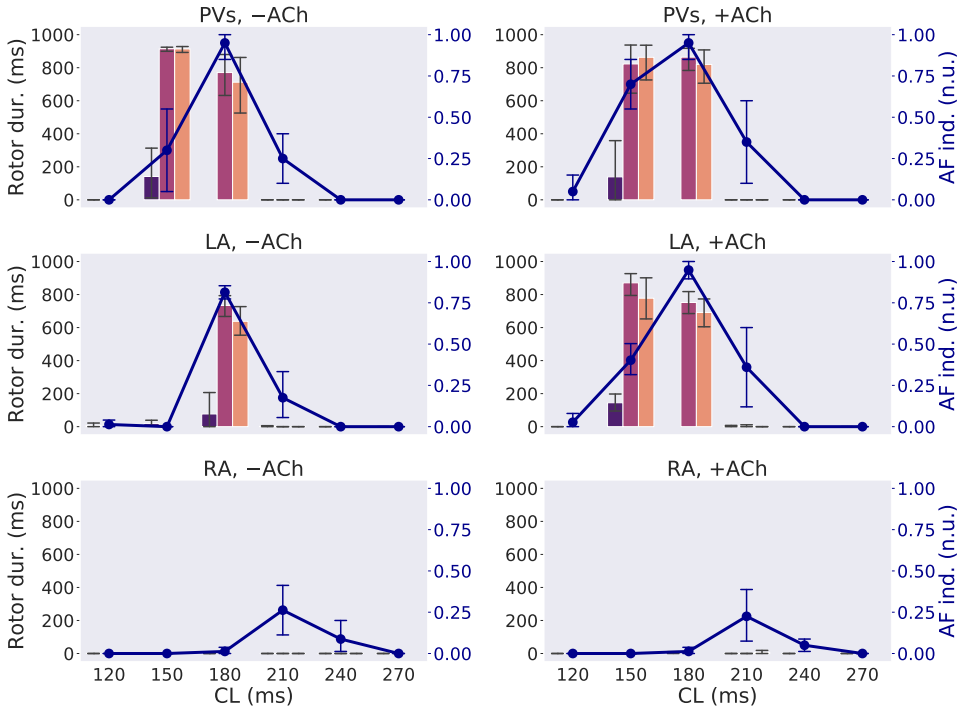


Fig. 3.14. Simulated ECG signals of (a) an IFS, (b) a DFS and (c) an RAF, corresponding to Fig.3.10(a), Fig.3.10(b), and Fig.3.12(a), respectively, after channel-wise normalization. a.u.: arbitrary unit.



(a) All rotors (PS duration $>$ 10 ms)



(b) Sustained rotors (PS duration $>$ 100 ms)

Fig. 3.15. Total duration ("Rotor dur." given by the bars) and AF inducibility ("AF ind.", given by the line plot) of (a) all rotors (PS duration $>$ 10 ms) and (b) sustained rotors (PS duration $>$ 100 ms), as functions of FS CL on LA, RA and PVs, with ("+") or without ("−") ACh. Vertical bars of the line plots show the 95% confidence interval of the mean of all patient meshes. CL groups accounting for less than 5% of FS in each panel were excluded from analysis and the bar charts. n.u.: normalized unit.

Table 3.4. The increase in AF inducibility from FS by adding ACh regulation, tested by a one-sided proportion z-test and grouped by FS CL and location. $+$ACh and $-$ACh denote with and without ACh, respectively.

FS CL (ms)	FS location	AF inducibility ($+$ACh, $-$ACh)	p-value ($+$ACh $>$ $-$ACh)
120	PV	0.05, 0	> 0.05
	LA	0.03, 0.01	> 0.05
	RA	0, 0	> 0.05
150	PV	0.70, 0.30	$<$ **0.01**
	LA	0.40, 0	$<$ **0.0001**
	RA	0, 0	> 0.05
180	PV	0.95, 0.95	> 0.05
	LA	0.95, 0.81	$<$ **0.01**
	RA	0.01, 0.01	> 0.05
210	PV	0.35, 0.25	> 0.05
	LA	0.36, 0.17	$<$ **0.01**
	RA	0.23, 0.26	> 0.05
240	PV	0, 0	N/A
	LA	0, 0	N/A
	RA	0.05, 0.09	> 0.05
270	PV	0, 0	N/A
	LA	0, 0	N/A
	RA	0, 0	N/A

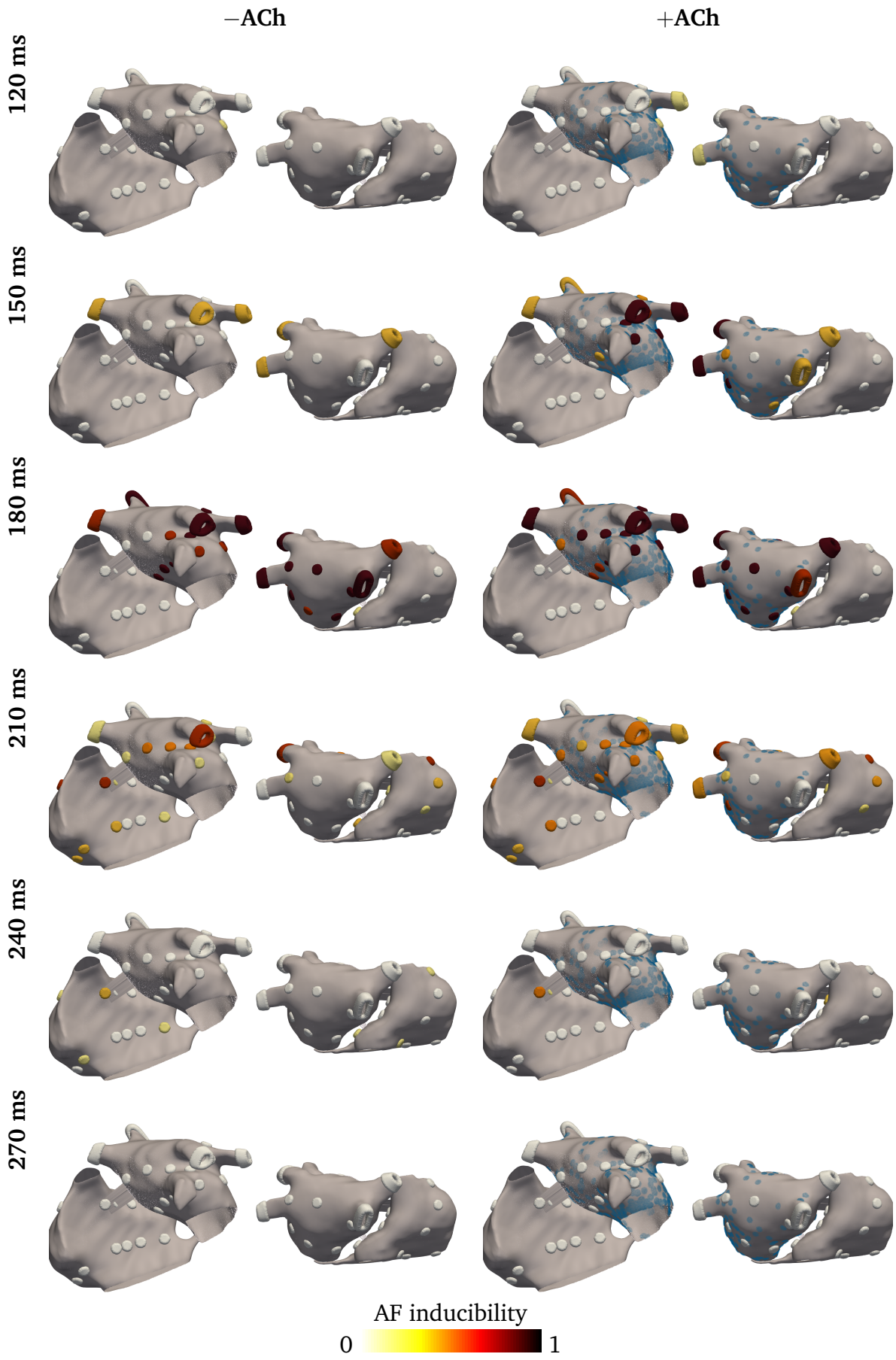
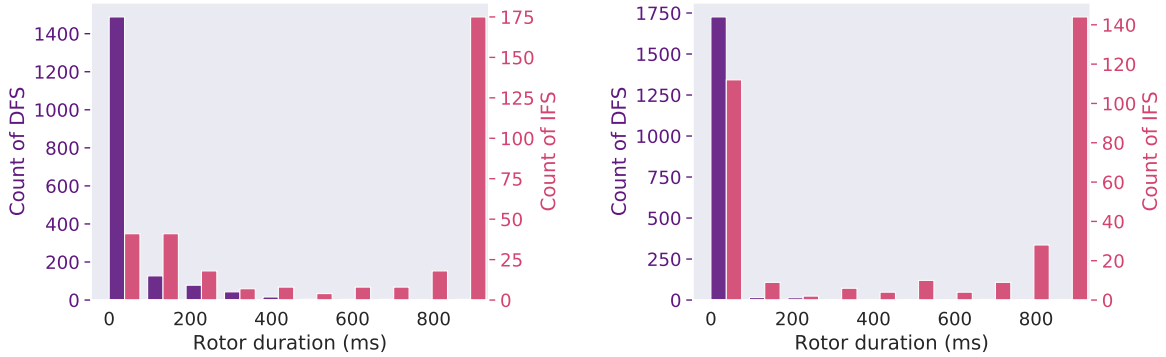


Fig. 3.16. AF inducibility for FS CLs ranging from 120 to 270ms across five patients, on the anterior and posterior view of a patient atria. The row shows the FS CL. Blue spherical markers are the locations of ACh islands. Markers of other colors represent focal size, where the color represents AF inducibility. Notation of "+" or "−" ACh means columns with and without ACh, respectively.

3.4.3 Distribution of rotor duration

Fig. 3.15 also shows that rotor duration was closely related to AF inducibility from FSs. For FSs with CLs of 150 – 180 ms from the LA and PVs, all cases where IFS had a longer duration of rotors or sustained rotors than DFS, were also associated with a high AF inducibility. The AF episodes induced by these IFS also had a similarly long rotor duration. On the contrary, short-lived rotors were common in CL of 210 ms, with a slightly higher average duration of all rotors presenting during DFS episodes than during IFS and FAF episodes, for all six conditions in Fig. 3.15(a).



(a) All rotors (PS duration $>$ 10 ms)

(b) Sustained rotors (PS duration $>$ 100 ms)

Fig. 3.17. Distribution of rotor duration for DFS and IFS using (a) all rotors (PS duration $>$ 10 ms) and (b) sustained rotors (PS duration $>$ 100 ms). In (a), the presence of a rotor in IFS (99.1%) is more frequent than in DFS (52.7%), with a p-value $<$ 0.0001 by one-sided proportional z-tests. In (b), the presence of a sustained rotor in IFS (65.8%) is more frequent than in DFS (2.6%), with a p-value $<$ 0.0001 by one-sided proportional z-tests.

For FSs with the presence of rotors, AF was more likely to be sustained than without. During FS-driven episodes, longer rotor duration was associated with higher AF inducibility, with an Area Under the Receiver Operating Characteristic Curve (AUROC) of 0.94 using the duration of all rotors (PS duration $>$ 10 ms) and an AUROC of 0.82 using the duration of sustained rotors (PS duration $>$ 100 ms) to separate between IFS and DFS.

To further explore this relationship, we calculated the distribution of rotor duration for DFS and IFS in Fig. 3.17 for both sustained rotors and all rotors. Most DFS were accompanied by short-lived rotors, and IFS were driven by macro-reentries, short-lived or sustained rotors, with a rotor duration of 0 – 100 ms for both all rotors and sustained rotors in Fig. 3.17. Rotors and sustained rotors appear more frequently in IFS than in DFS with a p-value $<$ 0.0001 by a one-sided proportional z-test. The remaining 34.2% of IFS without sustained rotor were mainly driven by short-lived rotors or macro-reentries, both of them were mostly induced with a FS CL of 210 ms, regardless of the focal site and [ACh] (Fig. 3.15(b)).

3.4.4 ERP reduction and APD heterogeneity

As can be seen from Fig 3.15, the AF-susceptible CL range varied between focal sites and [ACh]. To study the causes, the ERP was estimated by searching from long to short CL for the boundary CL where the 1:1 atrial response with the FS disappeared. As an example, a complete conduction block was formed around the FS with CL = 150 ms (Fig 3.18(b)) but not with CL = 180 ms (Fig 3.18(a)), the ERP for this focus was, thus, estimated between 150 and 180 ms.

For the majority of cases without ACh, the ERP was estimated as 150 – 180 ms for the LA FS, and 180 – 210 ms for the RA FS, longer than the LA FS. The AF-susceptible FS CLs were shorter in the LA than in the RA (in Fig 3.15), which was correlated with a shorter ERP in the LA than in the RA.

Likewise, the AF-susceptible FS CLs were shorter in the LA with ACh than without ACh, correlated with the shorter average tissue ERP for LA with ACh than without ACh. As an example, after adding ACh islands surrounding the LA focus in Fig 3.18(b), the unidirectional conduction blocks were seen around the focus, and a reentry was initiated (Fig 3.18(c)), which is equivalent to ERP shortening.

Furthermore, ACh regulation increased the spatial APD heterogeneity of the atrial tissue, which provided a substrate for reentry formation and maintenance. A higher variability in the APD on the LA tissue was also observed in the case with ACh in Fig 3.18(f), compared to without ACh in Fig 3.18(d,e).

3.5 Discussion

3.5.1 Rotor development during FS-driven episodes suggests an arrhythmogenic substrate

An important observation we made was that the AF sustainability was correlated with the presence of sustained rotors when the FS was active. This is based on the observation that IFS were predominantly associated with the presence of rotors. There was also a high AUROC score of 0.94 (all rotor duration) and 0.82 (sustained rotor duration) to predict the AF inducibility from FS. This observation provides a mechanistic basis for predicting the AF inducibility after the focal ablation using pre-ablation signals, where the computer model serves a deductive purpose.

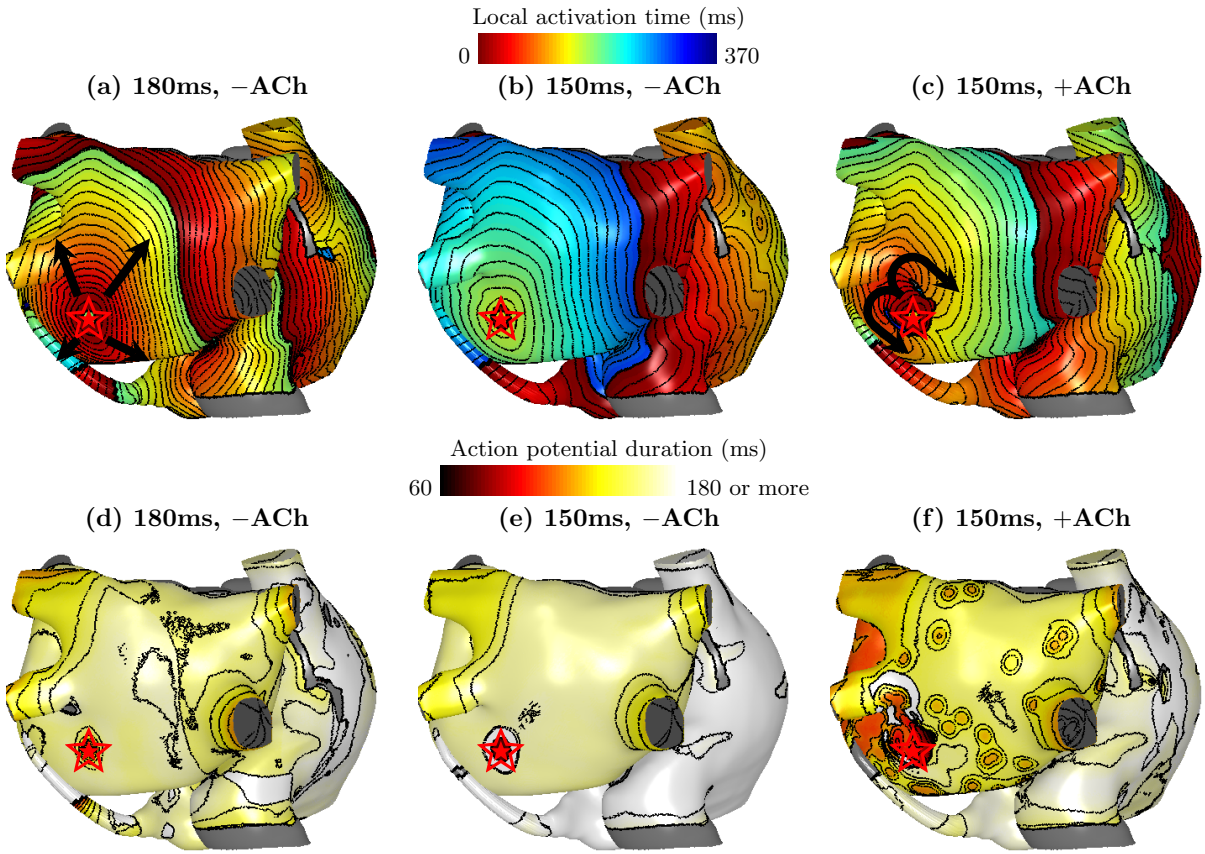


Fig. 3.18. Local activation time maps (a–c) and action potential duration maps (d–f) following the discharge of a FS (red stars) from the same site on the LA posterior wall $(\alpha_{LA} = 0.4, \beta_{LA} = 0.2)$**, on the atrial mesh of Patient 1, under focal cycle lengths of 150 ms and 180 ms, with ("+") and without ("−") ACh.** The directions of propagation from the FS were shown by the black arrows. A gray color indicates tissue that was not excited. Isochrone lines are drawn at an interval of 10 ms. The focal activation propagated in (a) and (c) but was blocked at (b). (f) shows the highest spatial heterogeneity in the action potential duration.

3.5.2 Factors that impact AF inducibility

According to the leading circle theory [190], the functional reentry was maintained in the form of the shortest circuit that has a length equal to the wavelength, defined as the product of average conduction velocity and average ERP. A shortening of wavelength due to ERP shortening and conduction slowing requires a shorter FS CL in order to initiate a reentry by FS. The AF inducibility from FS varied under different pacing and substrate conditions. The critical FS CL range associated with a large AF inducibility existed for all regions with ERP, as can be seen in Fig 3.15.

The difference of the critical CL range under varied regions and [ACh] was correlated with ERP shortening and spatial APD heterogeneity. For LA and PV foci, APD heterogeneity and the conduction slowing were found at the LA/PV junction [165], an important substrate for AF formation. For the RA FSs, APD heterogeneity existed transmurally between pectinate muscle and the RA epicardium [51], but they only produced short-lived rotors in our observations (Fig 3.15). ACh regulation further contributed to the spatial APD heterogeneity (Fig 3.18), consistent with other modeling studies [169, 170]. An increase in AF inducibility was observed at the area where the ACh islands were densely distributed (Fig. 3.16). Little difference in AF inducibility on the RA foci was found by adding ACh on the LA, which shows that it is the proximity to APD heterogeneous region that promotes AF inducibility.

Moreover, under the same atrial area, the maintenance of sustained AF requires reentries with a short wavelength [113], which impacts the AF inducibility in different regions and [ACh]. As the PV has the smallest ERP followed by LA, AF inducibility was also the highest for FS on the PVs, with the LA being the second highest in our results, which also agrees with non-invasive mapping results of Haïssaguerre and Rudy [41, 42, 40] and intracardiac rotor mapping results [191]. The ACh on the LA also increased the AF inducibility in LA and PVs.

3.5.3 Related studies

A tissue-level in-silico study of AF inducibility from FSs [115] showed that both focal timings and locations were essential to AF inducibility from FSs, with the spatial dispersion of APD as an underlying mechanism. We removed one of the timing factors, the coupling interval conduction between the sinus nodal stimulus and FS, by excluding the sinus rhythm from the experiments. This was based on the assumption that the low-frequency sinus rhythm plays a limited role in affecting the higher-frequency FS [183]. The consistency with previous findings suggested that our generated episodes represented patient AF faithfully. The smaller search space also reduced the required number of simulations needed for training robust prediction models. Moreover, we added ACh regulation in the experiments which further diversified our synthetic dataset in studying

factors that impact AF inducibility from FSs.

3.5.4 Computer models to enable supervised machine learning on BSPMs

As the ground truth of AF mechanisms is hard to obtain clinically, computer models which represented human atria with AF conditions were used as digital twins, to assist the development of diagnostic tools. For the purpose of improving the interpretation of BSPM for diagnosis, we focus on representing AF mechanisms on the organ level in order to simulate BSPMs. Meanwhile, a certain degree of realism is still required for modeling electrical propagation, including realistic conductivity and tissue heterogeneity. That is why we considered detailed atrial modeling with high-resolution patient-specific geometries. We showed that we were able to model functional and anatomical reentries, macro-reentries, and ACh effects. The critical FS CL range for high AF inducibility existed in all regions of LA, PVs and RA. In particular, from our model, the FS CL of 180 ms was associated with the highest AF inducibility, which is also very close to the average FS CL of 175 ms observed in AF patients[4], suggesting that our models were a good representative of the AF patient population.

It is known that the accuracy of the machine learning model increases with more data [192]. Therefore, we generated a total number of 2977 detailed AF simulations, amounting to more than 6 terabytes of storage, covering mechanisms ranging from organized to disorganized, including episodes driven by FS only, reentrant sources only, a mixture of reentrant and FS, as well as short-lived rotors (Fig. 3.9), with initial conditions of various locations and CLs over five atria with differences in anatomy and locations.

Another important purpose for building the computer model was to show the statistical correlation between the signals and prediction targets, required by a machine learning model. We showed two biophysical bases of using pre-operative signals to predict the effect of focal ablation: (1) a high AUROC of 0.94 between the rotor duration and AF inducibility from FS, and (2) CL of the FS is an important indicator of AF inducibility of FS. Both implied a causal relationship, stronger than a correlation relationship, between the pre-operative signals and the outcomes for the focal ablation. Both mechanistic insights also inspired us to look for methods to extract information of FS CL and rotors from the BSPMs, in order to infer the AF mechanisms non-invasively, the topic of the next chapter.

3.5.5 Limitation

The same model parameters were used across different meshes. They could have been varied to reflect a more diverse setting. We assumed our targeted paroxysmal AF patients were screened at an early stage of AF, so we did not add scar or fibrosis to our model, although this issue was mitigated by the introduction of ACh. Adding fibrosis could further enrich our dataset to represent a wider AF population. We simulated focal ablation by stopping focal discharges, without introducing ablation lesions, and allowed only 1000 ms to assess the AF sustainability after the ablation of the foci, which could be extended to a longer duration. We approximated BSPMs and ECGs with extra-cellular potentials on the vest electrodes, without inscribing the conductivity properties from the heart to the torso surface, which could still introduce small error. These limitations could be relaxed in future work.

Chapter 4

Non-invasive detection of focal source and arrhythmogenic substrate

4.1 Introduction

In the last chapter, we introduced the construction of a synthetic dataset by computer modeling. In this chapter[1], we present a method to classify AF source type, based on BSPM recordings of f-waves, using machine learning with second-order blind source separation (SO-BSS) to exploit common periodic structures among different signal channels.

The outline of this chapter is shown in Fig. 4.1. Recall that we denoted an FS as an ***Initiator-type FS (IFS)*** if AF was sustained after the focal ablation, and as a ***Driver-type FS (DFS)*** otherwise. AF after focal ablation for IFS was denoted as an ***FS-induced AF (FAF)***. A reentry driven AF episode was denoted as a ***Reentrant AF (RAF)***. AF sustainability after removing the FS indicated ***AF inducibility*** from the FS on the atria. We identified the classification of ***FS presence***, ***FS location*** and ***AF sustainability*** as targets of AF ablation, as shown in the motivation box of Fig 4.1. The presence of FS requires its hindrance or removal for treatment, such as a focal ablation procedure. The localization of foci could speed up their mapping. An arrhythmogenic substrate suggests that the atria tissue can support sustained AF, implying a need for area ablation, such as modifying the substrate [65] by disrupting the pathway of reentrant drivers [41], and/or limiting the available area for rotor movement [193].

To extract FSs from body surface potentials, we exploit their characteristics of high periodicity with a short CL [194]. However, high periodicity is not a unique charac-

[1]This chapter includes a part of the manuscript "Detection of focal source and arrhythmogenic substrate from body surface potentials to guide atrial fibrillation ablation" submitted to *PLOS Computational Biology*.

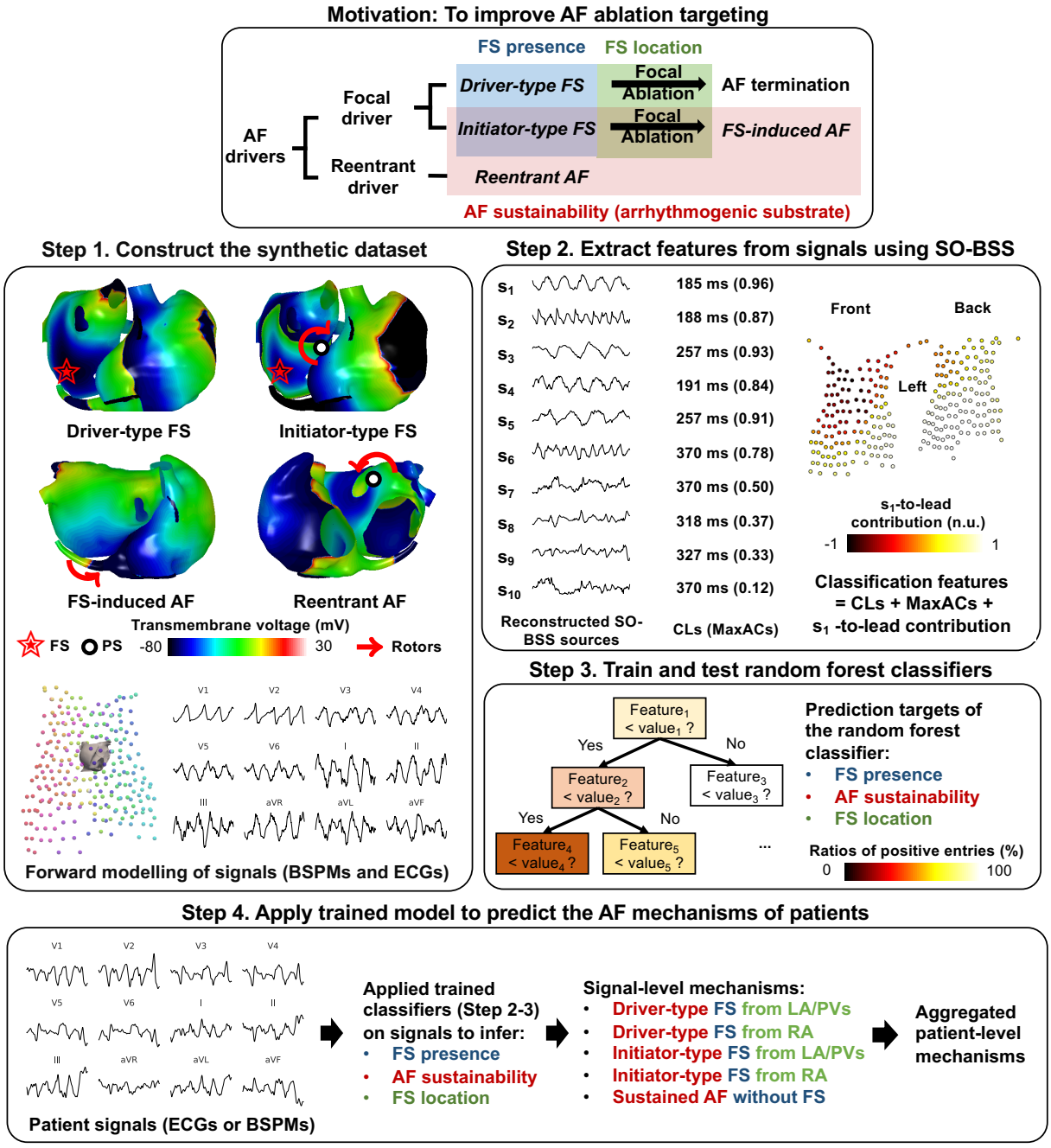


Fig. 4.1. An overview of our method. Steps 1 to 3 describe the training process of our classifiers on synthetic data, and Step 4 illustrates the application of the trained classifiers on patient signals to non-invasively detect their AF mechanisms. In Step 1 (the implementation was presented in Chapter 3), white spheres show phase singularity (PS) points, the red arrows show the wavefront movement of rotors, and the red stars mark the focal sites.

teristic in itself as, in theory, it is also exhibited by stationary rotors or reentries. It is therefore essential to perform multi-source analysis for all periodic sources.

SO-BSS is a class of well-established methods to extract periodic sources from signals by maximizing the periodicity of sources. It does not require *a priori* knowledge of the numbers of sources, which is usually a drawback for using blind source separation techniques in practice, such as FastICA [83]. The sources can be ranked according to their periodicity. Previously, SO-BSS methods have been used to extract atrial components from AF patient ECGs [143], to separate fetal and maternal contributions from maternal ECGs [195], as well as to extract FSs from AF patient intracardiac electrograms [155]. Furthermore, compared to traditional methods of analyzing periodic sources, such as the Fast Fourier Transform (FFT), the SO-BSS methods can extract periodic sources in relatively short signals with high resolution in the period or frequency estimated, which is more suitable for AF analysis.

We then adopted a nonlinear statistical model for high-dimensional features, a random forest classifier [86], to classify from multiple extracted sources. A random forest classifier is an ensemble of decision trees, where each node maximizes the information gain by splitting the parameter space of a feature. It then assembles these decision trees using bootstrap sampling, to reduce the risk of overfitting.

Given the hypothesis that sustained rotors during FS-driven episodes suggest an arrhythmogenic substrate that can be detected non-invasively with SO-BSS, the objective of this study was to non-invasively detect the presence of FSs and AF-susceptible substrates, as well as the location of any focal sites, in order to suggest appropriate ablation targets. We did so using computer simulations of realistic AF episodes driven by FS and/or reentrant sources, and by modeling focal ablation by removing FSs, to form digital twins in training the classifiers [19]. AF inducibility was ascertained by looking at AF sustainability after the FS was removed. From potentials computed on the torso, we sought to determine the prediction targets by extracting features using SO-BSS, and feeding these into a random forest classifier. This offers a novel algorithm to select appropriate ablation targets for AF patients, and detect the presence of malignant ectopic beats in the general population.

4.2 Methods

4.2.1 Equivalent atrial source extraction with SO-BSS

We first extracted K equivalent periodic atrial sources ($\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K$) and their CLs ($\mathrm{CL}_1$, CL_2, …, CL_K) from the body surface potentials, ranked by periodicity, using an SO-BSS method adapted from SOBI [81].

SO-BSS

Here we present atrial source extraction using an SO-BSS technique based on the algorithm SOBI [81]. The N-channel T-length body surface potentials, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]^T$ of shape $N \times T$, were assumed to have arisen from a linear mixture of K equivalent atrial sources $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K]^T$ of shape $K \times T$, and a linear relationship was assumed between the source and the signal. The signals are assumed to be a mixture of signals, with the following relationship,

$$\mathbf{X} = \mathbf{US}$$

where $\mathbf{U}$ is the unknown mixing matrix. The extraction of the source requires an estimation of its inverse matrix, $\mathbf{W}$.

The signals were first normalized channel-wise to allow a similar signal amplitude between different channels. Afterwards, K channels with the K largest eigenvalues through singular value decomposition on $\mathbf{X}$ were preserved, to reduce the computational burden in the following steps. The signal was then whitened to $\mathbf{X}_w = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_K]^T$ to remove correlation between channels, where K rows corresponding to the largest eigenvalues in Eq. (2.23) were preserved. The whitened signal was obtained by

$$\hat{\mathbf{X}} = \mathbf{D}_w^{-1/2}\mathbf{U}_w^T\mathbf{X} \tag{4.1}$$

where $\mathbf{U}_w$ is an orthogonal matrix of N eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N$, and $\mathbf{D}_w$ is a diagonal matrix with eigenvalues on the diagonal. This whitening process is also introduced in Section 2.3.2.

To estimate the second-order statistics of the signals, a whitened autocovariance matrix $C_x(\tau)$ with time-lag τ was defined.

$$C_x(\tau) = \hat{\mathbf{X}}_{\tau+1:T}\hat{\mathbf{X}}_{1:T-\tau}^T \tag{4.2}$$

where the subscripts of $\hat{\mathbf{X}}_{i:j}$ denotes that the times from i to j are included, and the subscripted T represents the length of the signal.

An iterative algorithm to perform joint diagonalization [81] was applied to calculate K joint eigenvectors $\{\tilde{\mathbf{w}}\}$ of length K over all $\{C_x(\tau)|\tau \in T_{FS}\}$, which constitute the rows of an estimated unmixing matrix $\tilde{\mathbf{W}}$ which approximate $\mathbf{W}$. The above process is also described in Section 2.3.2.

We ranked the sources from high to low according to the eigenvalue for each source, $\tilde{\lambda}_i$. The i-th predicted source $\tilde{\mathbf{s}}_i$ and its corresponding eigenvalue $\tilde{\lambda}_i$ obtained

from the joint diagonalization can be estimated as:

$$\tilde{\mathbf{s}}_i = \tilde{\mathbf{w}}_i^T \hat{\mathbf{X}}$$
$$\tilde{\lambda}_i = \max_{\tau \in T_{FS}} \tilde{\mathbf{w}}_i^T C_x(\tau) \tilde{\mathbf{w}}_i$$

where T_{FS} is a set of time delays.

Furthermore, the contribution of the i-th source $\mathbf{s}_i$ to i-th channel signal is given by

$$\mathbf{c}_i = \mathbf{U}_w \mathbf{D}_w^{1/2} \tilde{\mathbf{w}}_i \tag{4.3}$$

Extraction of machine learning features and dominant CL

The extracted source of SO-BSS may also be its own harmonics. To best estimate the true dominant CL, we set $T_{FS} = \{140, 141, \ldots, 370\}$ ms as the searched bound of source CLs. The upper bound of T_{FS} was set to cover all possible 2:1 conduction block conditions from FS with CL up to 185 ms, as the effective CLs of these FS on the surface signals were double the value of the focal CL. The lower bound marked the presumed minimal focal CL where the activation of FS could trigger 1:1 atrial responses.

The periodicity of the extracted source over a time lag τ was measured by an unbiased version of auto-correlation function (ACF), an empirical periodicity measurement for time series $\mathbf{s}(t)$ as a function of a time-lag τ [196]:

$$\text{ACF}(\tau) = \frac{T-1}{T-\tau-1} \frac{\sum_{t=1}^{T-\tau} (\mathbf{s}(t) - \bar{\mathbf{s}})(\mathbf{s}(t+\tau) - \bar{\mathbf{s}})}{\sum_{t=1}^{T} (\mathbf{s}(t) - \bar{\mathbf{s}})^2}$$

with $\bar{\mathbf{s}}$ being the expected value of $\mathbf{s}(t)$.

We denote $\text{ACF}_i(\tau)$ as the ACF function of the source $\mathbf{s}_i$ of time-lag τ. The estimated CL of each source $\mathbf{s}_i$ was adopted from the time-lag within a set of time lags T_{FS} resulting in the maximal auto-correlation score of the source $\mathbf{s}_i$,

$$\text{CL}_i = \underset{\tau}{\text{argmax}}\, \text{ACF}_i(\tau), \tau \in T_{FS} \tag{4.4}$$
$$\text{MaxAC}_i = \text{ACF}_i(\text{CL}_i) \tag{4.5}$$

To approximate the dominant rhythm in the atria, a dominant CL was taken by the first CL_k from $k = 1, 2, \ldots, K$ that satisfied $\text{ACF}_k(\text{CL}_k) > \text{CI}_k(\text{CL}_k)$, where $\text{CI}_k(\tau)$ denotes the range of 95% confidence interval for estimating $\text{ACF}_k(\tau)$.

On a Intel® Xeon® Gold 6140 processor with 36 cores, the time to extract SO-BSS features on a 1000ms ECG signal sampled at 1 kHz was around 0.2 second with $K = 1$,

around 0.3 second with $K = 10$, and around 2.4 seconds with $K = 20$.

4.2.2 Classification

Three binary classifiers were formed: (1) ***FS presence***, comprised of the union of IFS and DFS as the positive class, and the rest being negative; (2) ***AF sustainability***, to predict if there is an arrhythmogenic substrate where AF can be sustained in the atria without focal discharges from the FS, where the positive class consists of the union of FS-induced AF, RAF and IFS, and the negative class is DFS; (3) ***FS location***, trained and evaluated on the FS-driven episodes, to determine if the focal site is located on the RA (positive class), or from the LA or PVs (negative class). We also evaluated ***AF sustainability (with FS)***, which used the trained classifier from ***AF sustainability*** to distinguish between IFS (positive class) and DFS (negative class).

An illustration of features for each classifier is shown in Fig 4.1 Step 2. The classification of FS presence and AF sustainability requires a combination of multiple periodic sources to effectively represent AF. To encode the sources in a vector representation, we used two characteristics, CLs and MaxACs, given by Eq. (4.4) and Eq. (4.5), to represent each source, as the ACF of most extracted sources, especially those with a high periodicity, resembled sine waves with different periods. For K extracted atrial sources, $2 \times K$-dimensional characteristic features were used as input features for classification.

The features of the FS location classifier were the $\mathbf{s}_1$-to-lead contributions, forming an N-dimension vector specifying the contributions to each lead of the normalized signal from the reconstructed highest-ranked periodic source $\mathbf{s}_1$, given by Eq. (4.3). This encoded the spatial information of the first periodic source. As the SO-BSS sources are sign-agnostic, the signs of coefficients were fixed to one where a majority of coefficients were positive. The coefficients of all lead-to-$\mathbf{s}_1$ vectors were re-scaled to have the same maximum and minimum.

The imbalanced sample sizes of each class may result in a classifier biased towards the majority class. To overcome this problem, we applied an oversampling algorithm SMOTE [197] to make sure that all classes had the same number of training examples. This was then followed by a random forest classifier [86], which fits a non-linear boundary between the two classes, as shown in Fig 4.1 Step 3. It achieves this by splitting the parameter space of each feature in a way that maximizes the information gain. We fixed the number of decision trees in the random forest model at 200.

To train and evaluate a model on the five simulated patient datasets, we adopted a nested leave-one-patient-out cross-validation method, which reduced bias in estimating the true error of the classifier on small datasets [198]. There were two nested cross-validation loops, where the outer loop split the training/validation set and the one-patient test set. In the inner loop, the classifier was trained on the training set,

and selected the best hyperparameter resulting from the highest classification accuracy on the one-patient validation set. A model was then retrained with the best hyperparameter on the entire training/validation set, and evaluated on the test set as the performance of the classifier. The hyperparameter to select was the number of SO-BSS sources, K.

The SO-BSS algorithm was implemented in Python (version 3.6). SMOTE was implemented by imbalanced-learn Python package (version 0.6.2) [199]. The decision tree was implemented by the scikit-learn Python package (version 0.22.1) [200].

4.2.3 Application of trained classifiers on paroxysmal AF patients

To apply our trained classifier to predict patient AF mechanisms (Fig. 4.1 Step 4), data of 54 paroxysmal AF patients, including pre-operative BSPM signals from a 252-lead vest (Medtronic, MN, United States), as well as their clinical follow-up records, were exported from the Bordeaux Cardiology Hospital. The mapping protocol was approved by the institutional review board, and written consent was previously obtained from all patients. All anti-arrhythmic medications were stopped 48 hours prior to the catheter ablation, and diltiazem was given to patients before the procedure to slow the atrioventricular conduction. During the procedure, AF was first induced with burst pacing, and PV isolation was performed. After the isolation, if the sinus rhythm or atrial tachycardia was not restored, sites of AF drivers (foci and/or stable rotor cores) mapped by the catheter and the ECGi solution by CardioInsight, were ablated. For all patients, AF was terminated at the end of all procedures.

BSPM signals were acquired at a sampling rate of 1 kHz. QRS complexes were extracted by the Pan-Tompkins algorithm [201]. T-wave were extracted automatically by identifying the segment with a large variance over the signals of all channels on the segment between two QRS complexes. Only f-wave segments longer than 800 ms were included for analysis. We excluded three patients under 30 years old, as well as one patient who had only one f-wave segment recorded. 13 ± 4 segments were analyzed for each patient. All signals were first filtered by a standard 2 – 30 Hz second-order Butterworth filter to remove high-frequency noise and baseline wandering, and then were normalized channel-wise. We computed the signal-level mechanism for each signal as in Table 4.1. The patient-level mechanisms were obtained by aggregating all signal-level mechanisms of each patient.

Table 4.1. Computation of signal-level mechanisms from the positive ("+") or negative ("−") prediction outputs of our classifiers. Note that the FS location classifier outputs whether the FS is on RA.

Mechanism	Classifier outputs		
	FS presence	AF sustainability	FS location
IFS from LA/PVs	+	+	−
IFS from RA	+	+	+
DFS from LA/PVs	+	−	−
DFS from RA	+	−	+
Sustained AF without FS	−	+	+/−

4.3 Results

4.3.1 Atrial source extraction with SO-BSS

To illustrate the advantages of SO-BSS, we show examples of sources decomposed from BSPMs of Patient 2 with SO-BSS, in IFS (Fig. 4.2), DFS (Fig. 4.3), and RAF (Fig. 4.4), with ECGs shown in Fig. 3.14. The ECG morphological differences are indistinguishable to the human eye between IFS (Fig. 3.14(a)) and RAF (Fig. 3.14(c)). FFT analysis also had a limited frequency resolution (1 Hz) on a 1000 ms 1 kHz signal, where IFS and RAF all showed the same DF of 4 Hz on the V1 lead. Their V1 signals also showed similar maximal values of ACF (0.68 and 0.67).

However, through extracting multi-source CLs, different degrees of CL alignment between SO-BSS sources were shown across different mechanisms. As the DFS contained only one type of periodic source, CLs of extracted sources were the most uniform in the DFS in nine out of ten sources. The sources of an IFS were less synchronized, with six ($\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_4$, $\mathbf{s}_6$, $\mathbf{s}_7$ and $\mathbf{s}_{10}$) that were synchronized (less than ± 10 ms). The RAF, however, had only three synchronized sources ($\mathbf{s}_1$, $\mathbf{s}_3$ and $\mathbf{s}_6$) with CLs similar to the dominant CL of 191 ms and its harmonics. This showed the necessity of combining multi-lead features to classify different AF mechanisms.

4.3.2 Spatiotemporal feature extracted by SO-BSS

To illustrate the usage of SO-BSS in extracting the latent spatiotemporal structure of the signals, we show the waveforms of the signals and estimated sources for the first 300 ms in Fig. 4.5, Fig. 4.6 for the three cases of DFS, IFS and RAF analysed above, where the signal length of 300 ms was chosen to cover at least one cycle of the signals and the reconstructed sources. Here, the contributions were continuous in adjacent leads in the Euclidean space, estimated from the covariance of signals, without any explicitly

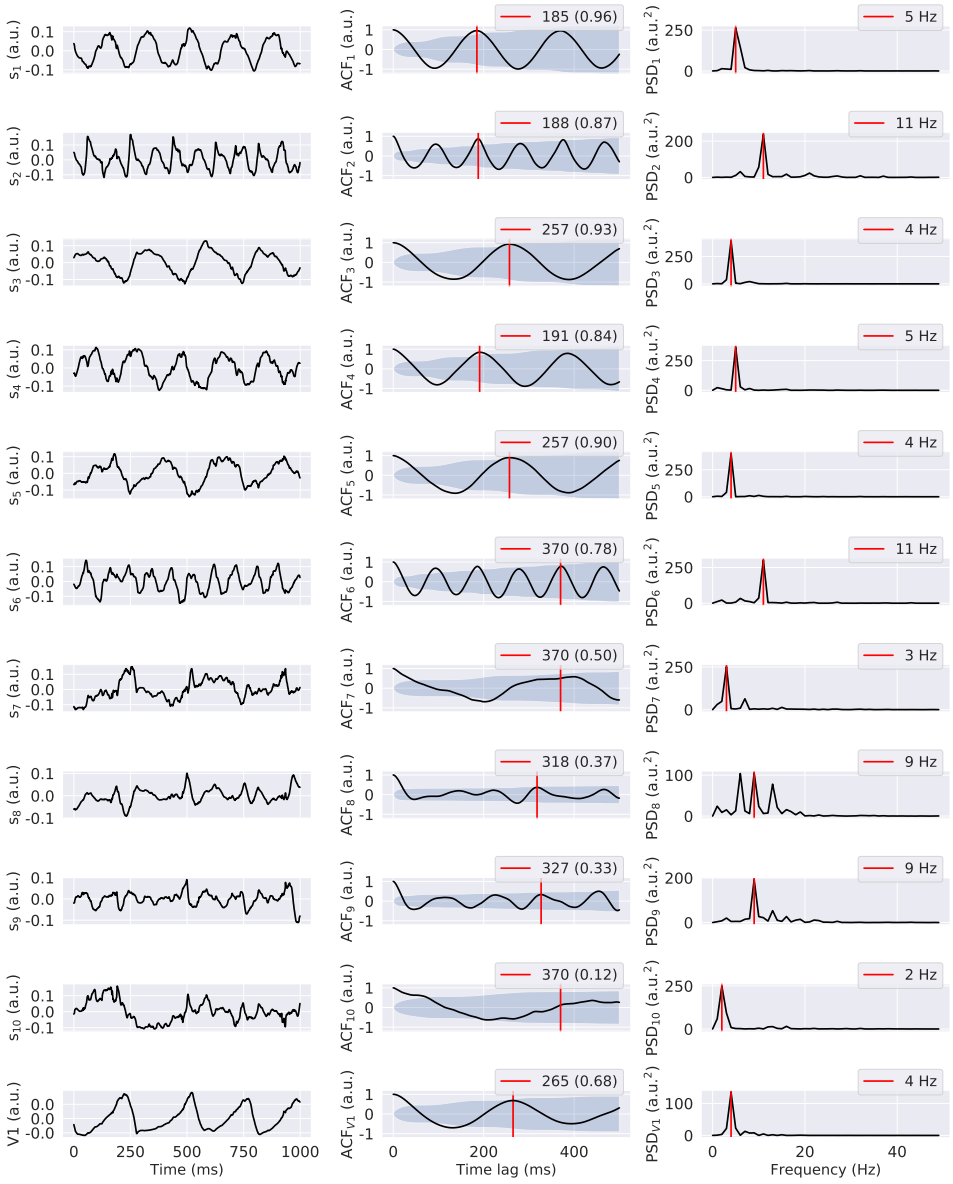


Fig. 4.2. V1-lead ECG signal and SO-BSS sources extracted from the BSPM of an IFS. Signals are shown for an IFS in patient 2 with a CL of 180 ms (the same as Fig 3.10(a)) at an LA focal site ($\alpha_{LA} = 0.2, \beta_{LA} = 0.2$) using $K = 10$. The first 10 rows show the top 10 sources, $\mathbf{s}_i$, ranked by their eigenvalues, respectively, with the bottom row showing the V1-lead ECG signal. The first column shows the signal amplitude over time. The second column shows the value of ACF($\mathbf{s}_i$) over time-lags up to 500 ms with shading showing the 95% confidence interval. The red bars indicate CL_i. Labels denote CL_i (MaxAC_i) for the top 10 rows, and CL_{V1} (MaxAC_{V1}). The third column shows the FFT power spectral density of the signal in the first column with the DF indicated by a red bar. a.u.: arbitrary unit.

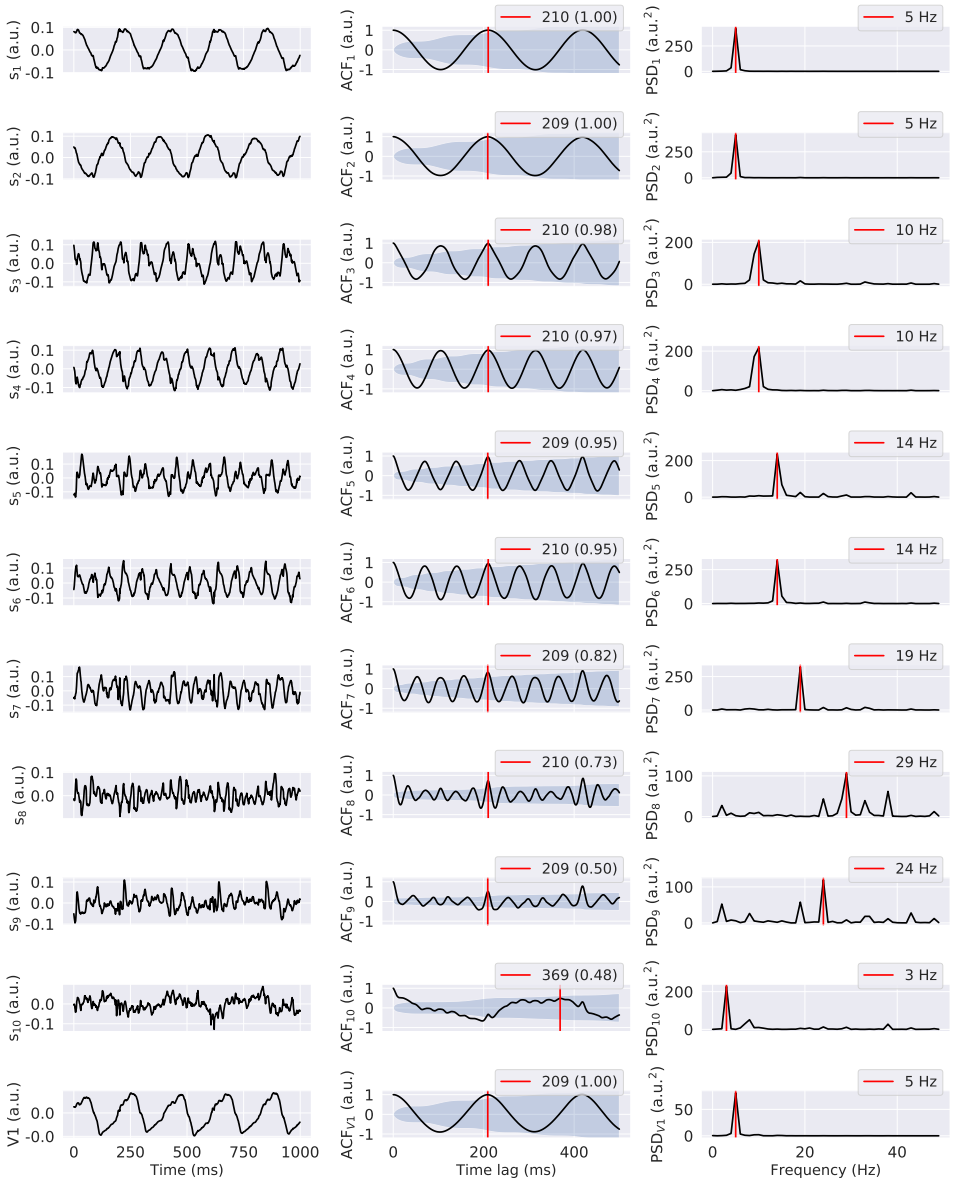


Fig. 4.3. V1-lead ECG signal and SO-BSS sources extracted from the BSPM of a DFS. Signals are shown for a DFS with a CL of 210 ms at an LA focal site ($\alpha_{LA} = 0.2, \beta_{LA} = 0.2$) with $K = 10$ for Patient 2 (the same as the case of Fig 3.10(b)). The legend is the same as Fig. 4.2.

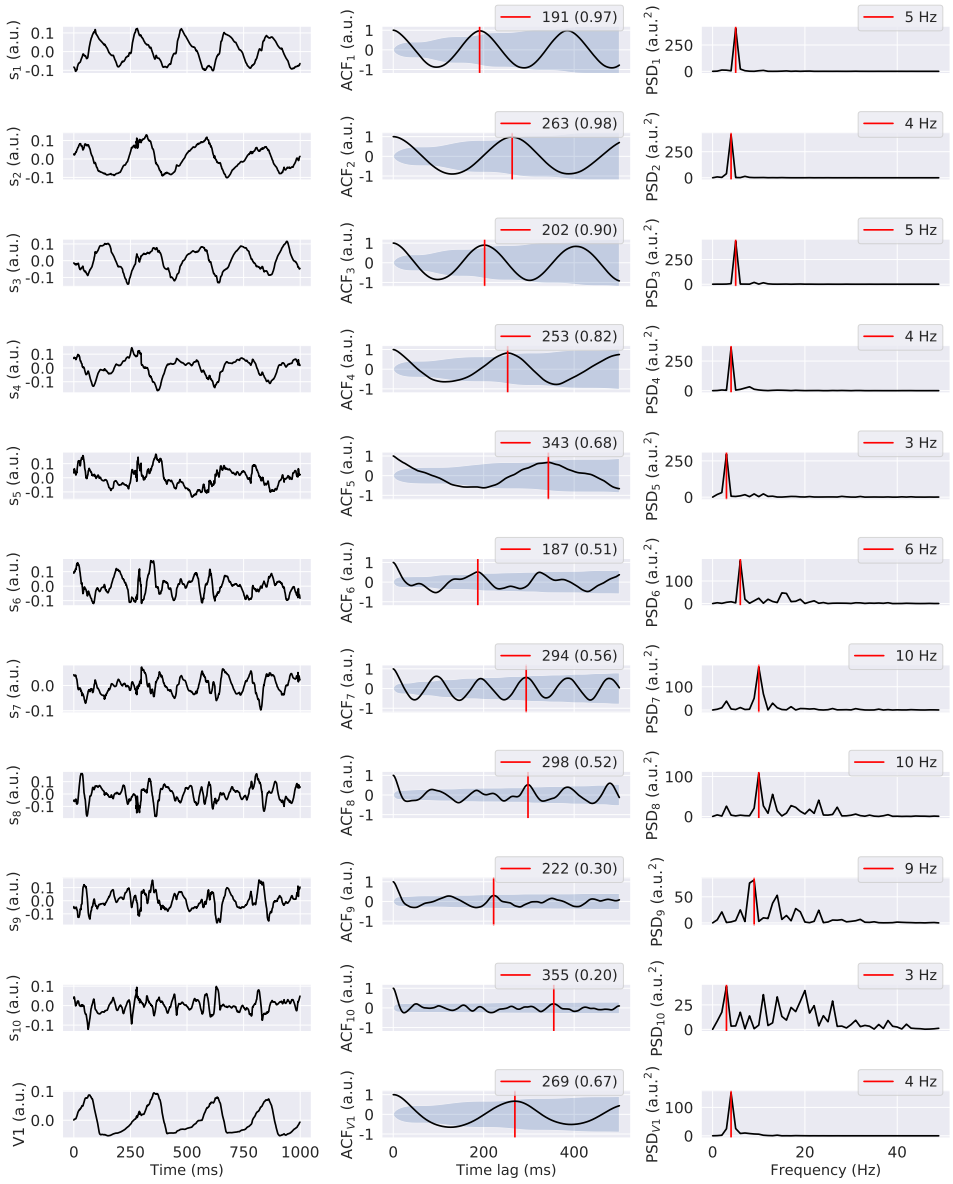


Fig. 4.4. V1-lead ECG signal and SO-BSS sources extracted from the BSPM of an RAF. Signals are shown for an RAF around the LA site at $(\alpha_{LA} = 0.2, \beta_{LA} = 0.2)$, without a central inexcitable region, were extracted from the BSPM with $K = 10$ on the mesh of Patient 2, the same as the case of Fig. 3.12(a). The legend is the same as Fig. 4.2.

encoded spatial information of the lead position. The two figures together show that the source extraction does not only cope with the simple case of DFS, but can also cope with multiple types of atrial activities. One can also verify that the signals of the BSPMs were contributed by the estimated sources shown at the right bottom, for all three types of simulations.

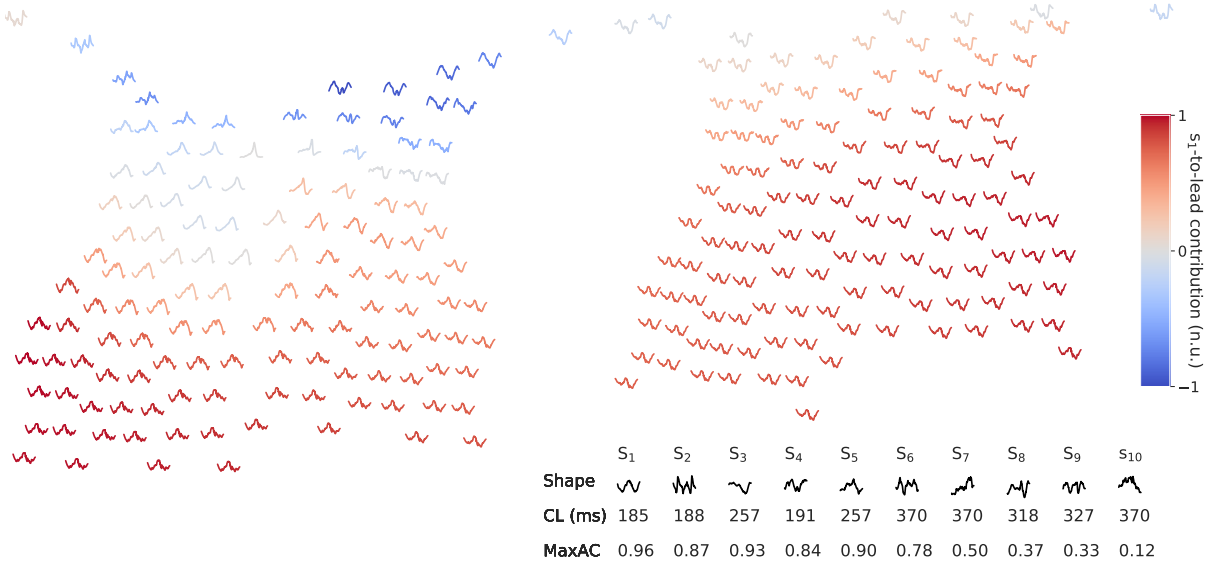
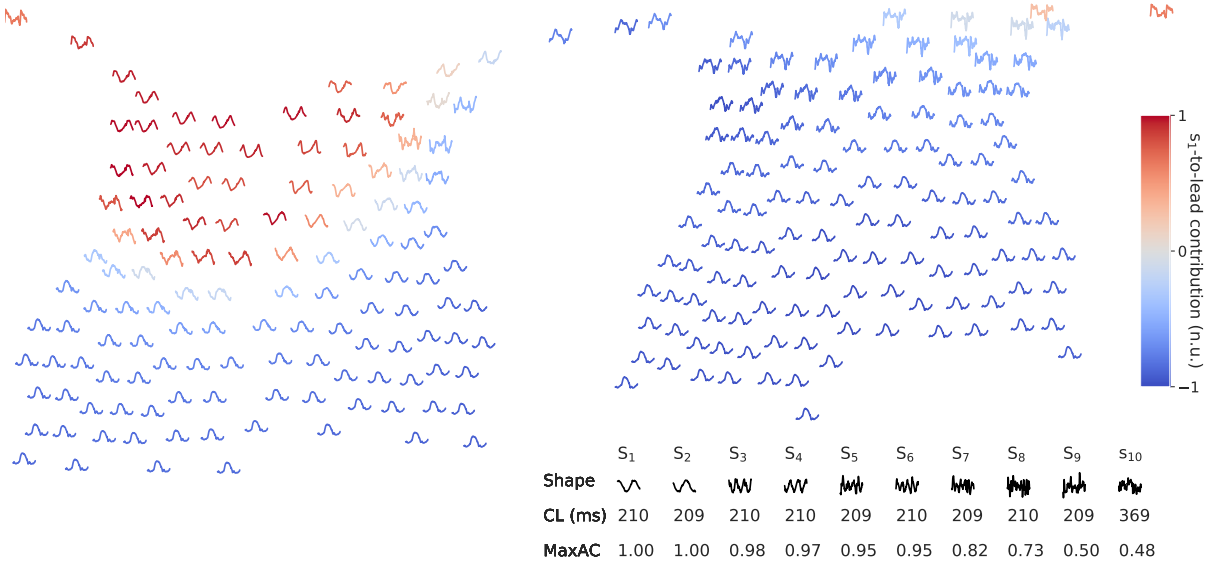


Fig. 4.5. BSPM signals and estimated sources extracted by SO-BSS ($K = 10$) after channel-wise normalization of the IFS in Fig. 4.2, up to 300 ms. The BSPM signals are colored by the rescaled value of s_1-to-lead contribution. The signals were plotted on electrode positions, where the unshown x-axis and y-axis are the rotational angle and the height respectively, in the same way as the vest electrodes shown in Fig. 4.1 Step 2. The s_1-to-lead contribution was linearly rescaled to a value between -1 and 1.

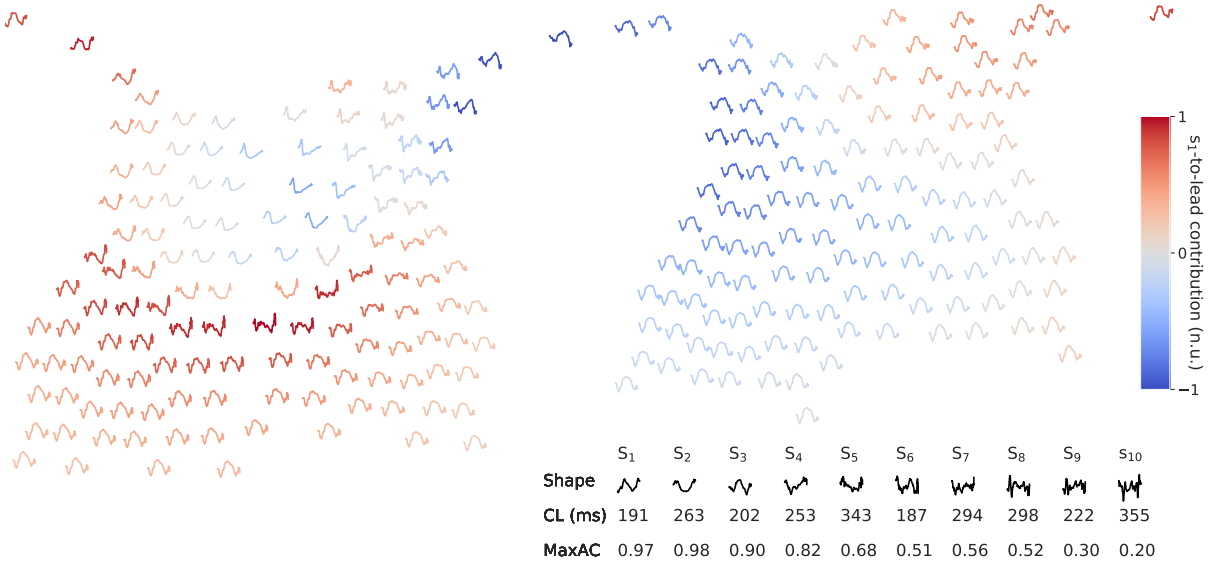
We also compared the spatial distributions of s_1-to-source contribution for FS with different CLs from the same UAC coordinate in Fig. 4.7, and spatial patterns are consistent across different FS CLs. Comparing the distributions between before and after the introduction of ACh on the same patient atria in Fig. 4.7 (a-b), the spatial patterns are similar, demonstrating a robustness against spatial heiterogenity in the atria. The spatial pattern is also insensitive to different patients (Fig. 4.7 (a) and Fig. 4.8(a)). However, it is sensitive to the change of focal site (Fig. 4.7 (a) and Fig. 4.8(b)). On the other hand, the IFS cases in these two figures altered the spatial patterns in each condition, because of the presence of reentrant sources.

4.3.3 Evaluation of classifiers

We evaluated the classification results of our features derived by SO-BSS with the following AF-complexity metrics in Table 4.2. The first metric was the ratio of FFT power spectrum density of 0–2×DF over 0–50Hz of the V1 lead, called AFFTr$_{2DF}$, which was



(a) The DFS in Fig. 4.3



(b) The RAF in Fig. 4.4

Fig. 4.6. BSPM signals and estimated sources extracted by SO-BSS ($K = 10$) after channel-wise normalization of (a) the DFS in Fig. 4.3, and (b) the RAF in Fig. 4.4, up to 300 ms. The BSPM signals are colored by the rescaled value of s_1-to-lead contribution. The legend is the same as Fig. 4.5.

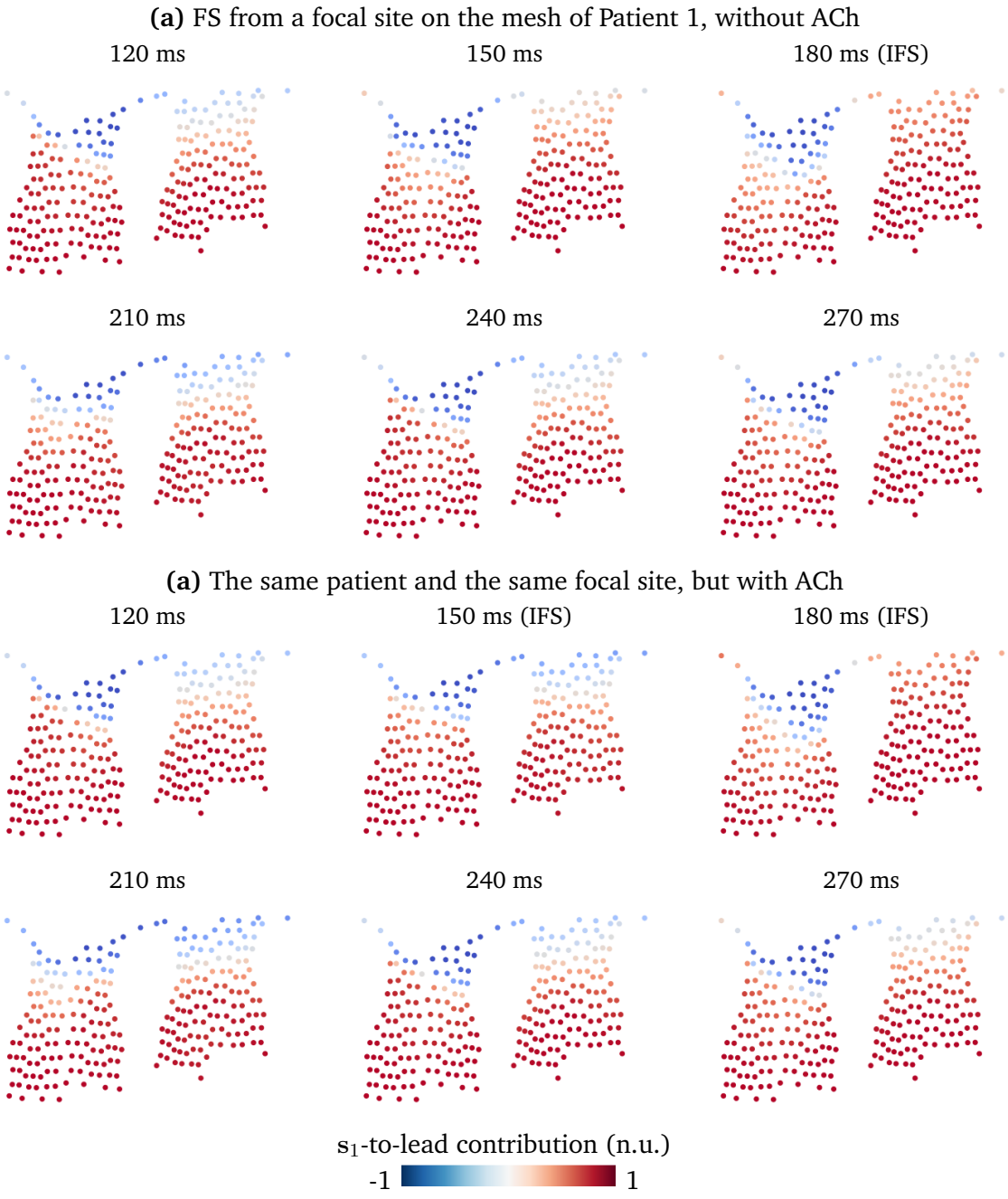


Fig. 4.7. Spatial patterns of s_1-to-lead contribution extracted by SO-BSS ($K = 1$) on an FS with CLs of 120 – 270 ms from a focal site of $(\alpha_{LA} = 0.2, \beta_{LA} = 0.2)$, which show insensitivity against different FS CLs, to the perturbation of ACh regulation (a,b). IFS may bring some perturbation to the patterns due to the additional reentrant sources in IFS. The electrode positions are plotted in the same way as the vest electrodes shown in Fig. 4.1 Step 2. All entries are DFS except for those which are marked as IFS.

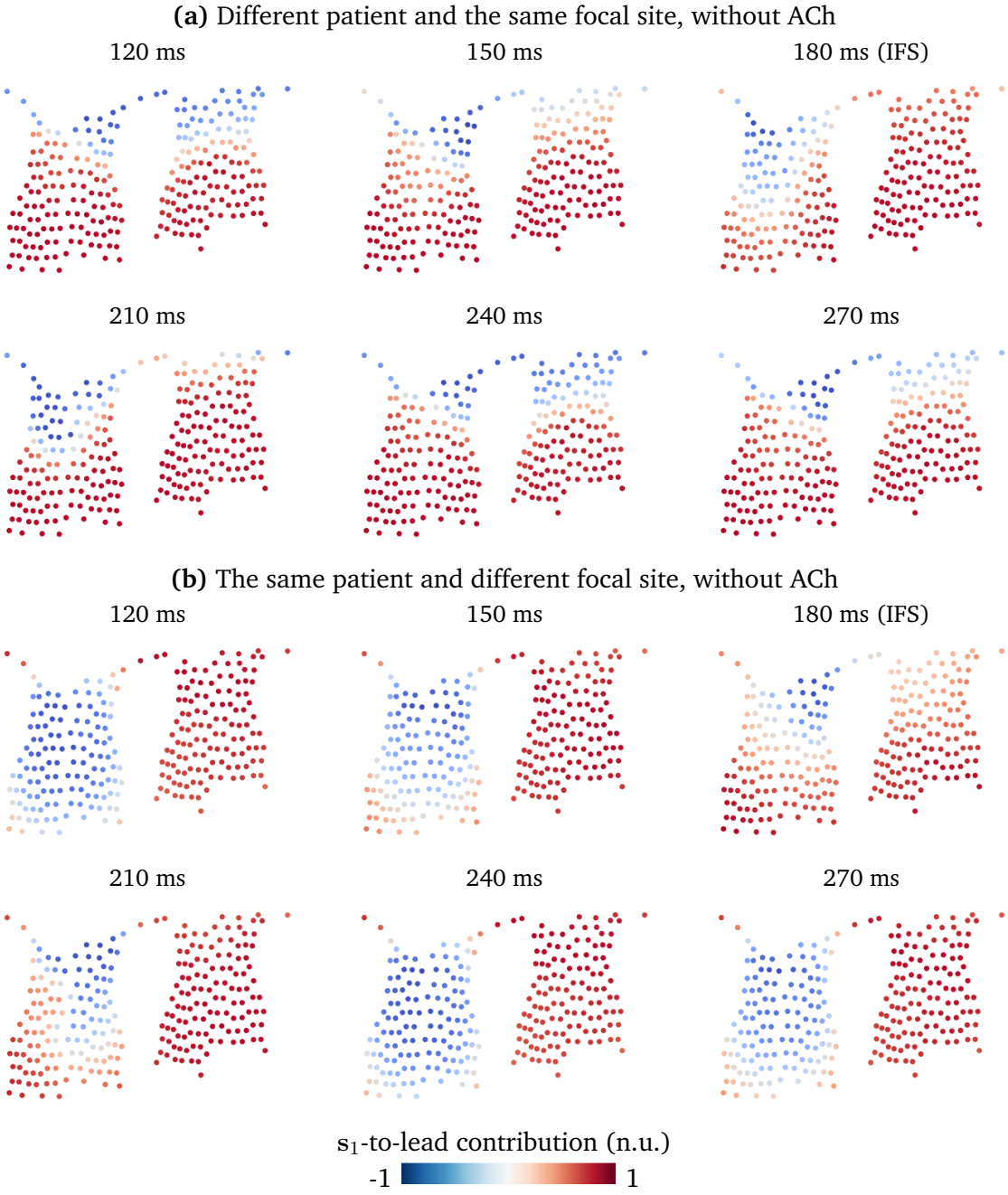


Fig. 4.8. Spatial patterns of s_1-to-source contribution extracted by SO-BSS ($K = 1$) on FS with CLs of 120 – 270 ms (a) from the same location of a different patient mesh with Fig. 4.7, and (b) from a different site $(\alpha_{LA} = 0.2, \beta_{LA} = 0.8)$ of the same patient. Compared with Fig. 4.7(a), this figure shows the spatial pattern remains relatively consistent to inter-patient variability (a), but sensitive to the change for FS originating from a different location (b). All entries are DFS except for those which are marked as IFS.

developed to distinguish between FS and reentrant sources [131]. The second metric was the nondipolar component index (NDI), which was used to predict the acute ablation outcome in persistent AF patients from BSPMs [15], measured as the proportion of the residuals outside of the three principal axes. We replaced the SO-BSS features with these metrics, and reused the same classification pipeline for training and testing, i.e., SMOTE and a random forest classifier.

Table 4.2. Testing scores using nested leave-one-patient-out cross-validation on all classification tasks over five patients, using features of SO-BSS, AFFTr$_{2DF}$ and NDI. Bold fonts mark the highest scores and the corresponding features.

Task	Feature (with signal used)	Accuracy (mean±s.d.%)	Precision (mean±s.d.%)	Recall (mean±s.d.%)
FS presence	**SO-BSS (BSPM)**	**90.7±1.2**	**96.8±1.0**	**89.9±0.8**
	SO-BSS (ECG)	**90.7±0.7**	**97.5±0.5**	**89.2±0.8**
	AFFTr$_{2DF}$ (V1)	57.8 ±1.3	73.8±0.8	62.3±2.4
	NDI (BSPM)	59.7±1.5	74.8±2.2	64.8±4.1
AF sustainability	**SO-BSS (BSPM)**	**92.8±1.5**	**90.0±3.0**	**92.7±3.6**
	SO-BSS (ECG)	**93.7±1.2**	**91.5±3.0**	**93.0±3.3**
	AFFTr$_{2DF}$ (V1)	58.3±2.3	48.3±3.6	53.1±3.5
	NDI (BSPM)	61.7±3.0	52.4±5.7	57.0±6.0
AF sustainability (evaluated on FS-driven episodes)	**SO-BSS (BSPM)**	**90.6±2.8**	**68.1±7.6**	**80.3±15.0**
	SO-BSS (ECG)	**91.8±1.9**	**72.2±7.3**	**81.1±14.3**
	AFFTr$_{2DF}$ (V1)	60.7±3.5	20.9±2.3	55.7±3.8
	NDI (BSPM)	63.6±4.2	23.3 ±4.9	57.1±5.8
FS location	**SO-BSS (BSPM)**	**80.0±6.6**	**81.2±9.0**	**75.0±8.4**
	SO-BSS (ECG)	61.0±5.2	59.3±6.8	53.6±8.4

On FS presence and AF sustainability, our method achieved high scores of accuracy, precision and recall, all close to or above 90% with variance $< 4\%$, significantly outperforming AFFTr$_{2DF}$ and NDI, showing the efficacy of our algorithm, as well as robustness to inter-patient variability in atrial geometries. When the AF sustainability classifier was evaluated on FS-driven episodes, the accuracy was still high, but the precision and recall scores slightly decreased. These scores were mainly impacted by FS with a CL of 210 ms, as shown in Table 4.3, which was expected given that the IFS had a different mechanism at CL of 210 ms as discussed in the previous section, and AF inducibility at this CL also presented a considerably larger inter-patient variation (12.5%) compared to other CLs ($\leq$ 3.2%). Although we did not apply any specific location information of the atria nor the torso leads, the estimation of the focal site region achieved moderately high accuracy of 80.0% and 61.0% from BSPMs and ECGs.

To understand how different mapping systems and numbers of sources K affect the classification performance, we calculated the mean and standard deviation of the accuracy score across all five outer-validation sets in Fig 4.9. Both FS presence and AF sustainability classifiers presented similar trends, where the accuracy increased up to a certain K ($K = 6$ for FS presence and $K = 4$ for AF sustainability) with little change afterwards, likely due to the inherently small number of sources with distinctive CLs. For the FS location classifier, with a larger K, the accuracy dropped for BSPMs but

increased for ECGs.

Table 4.3. Atrial fibrillation (AF) inducibility (mean±s.d.%), and testing accuracy scores for different methods (mean±s.d.%). Results are shown for different CLs (ms). Leave-one-out cross validation was used on the AF sustainability prediction incorporating all FS episodes, grouped by focal CL (ms). Bold fonts mark the highest scores. Signals used are given in parentheses. SO-BSS took $K = 10$ sources as an input parameter.

CL	AF inducibility	**SO-BSS (BSPM)**	AFFT_{2DF} (V1 ECG)	NDI (BSPM)
120	1.9±0.8	**98.8±1.2**	61.1±8.1	66.2±8.6
150	14.3±2.1	**88.3±3.1**	71.4±4.7	64.9±9.1
180	49.1±2.2	**96.9±3.2**	64.3±6.1	70.8±3.5
210	26.0±18.5	**64.8±12.5**	50.0±6.2	52.5±17.2
240	5.2±3.0	**98.3±1.9**	62.0±7.4	66.8±5.9
270	0	**98.9±2.6**	55.1±5.3	60.2±6.0

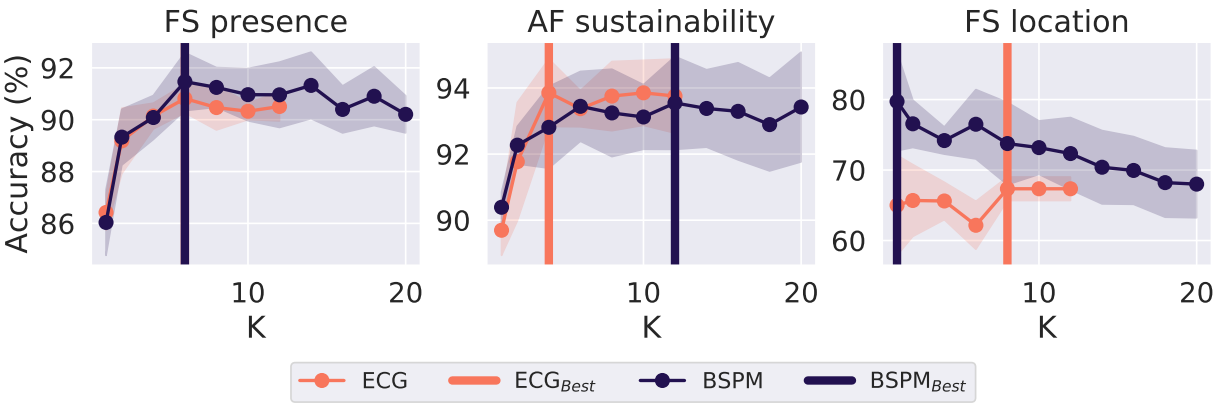


Fig. 4.9. Classifier performance as a function of K. Means (points) and standard deviations (shaded areas) of the test accuracy scores using leave-one-patient-out cross validation are shown for changing K and lead systems (ECGs or BSPMs), on the classification tasks, with the best K for the highest accuracy (ECG_{best} or BSPM_{best}) in each setting shown by vertical bars.

Table 4.4. Evaluation of heart position on classifier performance. Dominant CL estimation, FS presence, and AF sustainability classifiers, with 5 and 10 cm translations, and 10 and 20 degree in rotation, on the dataset of Patient 2, using $K = 10$ and BSPM. Measures were computed using a 252-lead vest. Absolute dominant CL difference refers to the absolute difference in the dominant CLs between the original setting and the variation, for which the mean and the standard deviation (s.d.) are shown. The asterisk ($*$) marks the situation when a heart transforms out of the boundary of the vest and the results were excluded from the analysis of mean and standard deviation, and the maximal difference.

		Original	Rotate (degree)				Translate -10cm			Translate -5cm			Translate +5cm			Translate +10cm		
			-20	-10	10	20	x	y*	z	x	y	z	x	y*	z	x*	y*	z*
Absolute dominant CL difference	Mean (ms)	0.0	3.05	2.41	3.39	4.35	2.99	2.53*	3.76	3.14	4.65	2.27	4.78	7.09*	2.92	2.06*	1.21*	2.91*
	s.d. (ms)	0.0	20.15	17.35	20.57	24.18	17.49	17.26*	22.21	20.15	22.78	15.44	23.83	30.02*	18.94	16.14*	11.49*	18.48*
FS presence	Accuracy (%)	91.3	91.0	91.7	91.0	91.0	91.0	89.4*	91.3	91.2	91.2	91.5	91.7	90.5*	91.7	90.7*	91.3*	91.2*
	Precision (%)	97.9	97.6	97.9	96.9	97.4	97.6	91.5*	97.7	97.9	97.4	97.4	97.9	96.4*	97.4	98.1*	93.8*	97.4*
	Recall (%)	89.0	88.8	89.5	89.5	89.0	88.8	92.9*	89.3	88.8	89.3	89.8	89.5	89.3*	90.0	87.9*	93.3*	89.3*
AF sustainability	Accuracy (%)	92.5	91.8	92.3	91.7	92.0	92.6	92.9*	92.6	92.6	92.5	92.0	92.1	92.3*	92.9	92.3*	91.5*	92.5*
	Precision (%)	92.1	91.3	92.0	90.4	90.8	93.3	94.7*	91.8	92.1	92.1	91.7	91.1	92.6*	92.1	91.1*	92.8*	93.3*
	Recall (%)	91.1	90.4	90.7	91.1	91.4	90.0	89.3*	91.8	91.4	91.1	90.4	91.4	90.0*	92.1	91.8*	87.9*	89.6*
AF sustainability (from FS)	Accuracy (%)	88.8	87.9	88.6	87.9	88.3	89.3	89.5*	89.0	89.0	89.0	88.1	88.3	88.6*	89.5	88.8*	87.9*	89.0*
	Precision (%)	69.9	67.1	69.4	65.8	67.1	73.1	76.7*	69.7	70.3	70.3	68.1	67.5	70.6*	71.1	68.4*	69.8*	72.7*
	Recall (%)	67.1	64.5	65.8	68.4	69.7	64.5	60.5*	69.7	68.4	68.4	64.5	68.4	63.2*	71.1	71.1*	57.9*	63.2*

To test the sensitivity to different vest placements, we varied the vest positions by 10 and 20 degrees around the z-axis pointing from the centre of mass of the vest electrodes, and 5 and 10 cm in the x, y, and z directions using the same coordinate system as in Fig. 3.4, on the dataset of Patient 2, with $K = 10$ sources for SO-BSS with BSPMs as shown in Fig 4.10. These variations are considered extreme cases for intra-patient variations, only designed to test the robustness of our algorithm. The results of the sensitivity tests for the FS presence, AF sustainability classifiers and the dominant CL estimation are shown in Table 4.4. The absolute differences of dominant CLs between all vest variations and the original vest placement were small, with a mean ≤ 4.78 ms and standard deviation ≤ 24.8 ms. All vest placements received similar classification scores with respect to each other, with the maximal absolute difference being 4.1%.

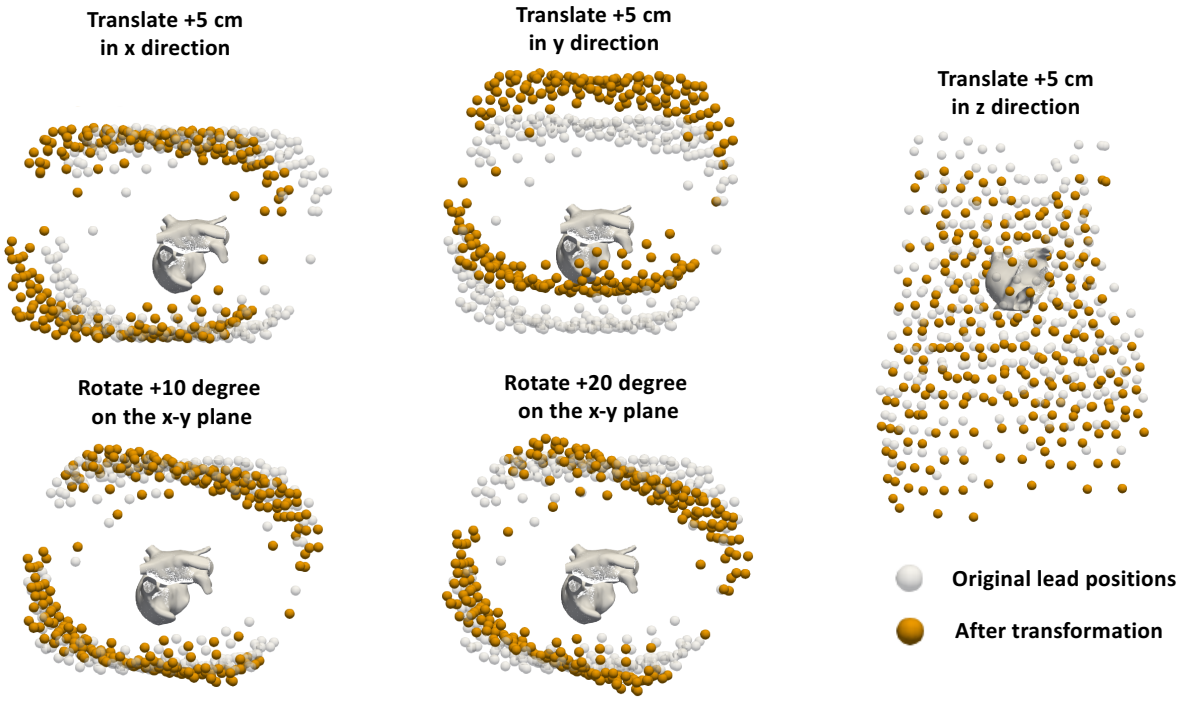


Fig. 4.10. Examples showing the electrode placement after translation (5 cm) and rotation (10, 20 degrees) of vest leads, using the same coordinate system as in Fig. 3.4.

4.3.4 Feature interpretation of classification

MaxAC values, as part of the feature vector, encoded information about multi-source periodicity as signatures of atrial events. The MaxAC values of sources extracted by SO-BSS with $K = 10$ from BSPMs were visualized by being projected on to two-dimensional orthogonal bases using acPCA [16], colored by their ground-truth mechanisms or CLs, as shown in Fig 4.11 (a) and (b), respectively. The PCA captured variances $93.2\pm0.7\%$ and $93.2\pm0.5\%$ in the reduced dimensions. On all five patients, the DFS could be easily separated from the other mechanisms, but there was a more complex decision boundary

between IFS and DFS. CL clusters could also be seen in the feature space, such as CL clusters of 150 ms and 180 ms, both of which were AF-susceptible CLs, were adjacent to each other in all patients.

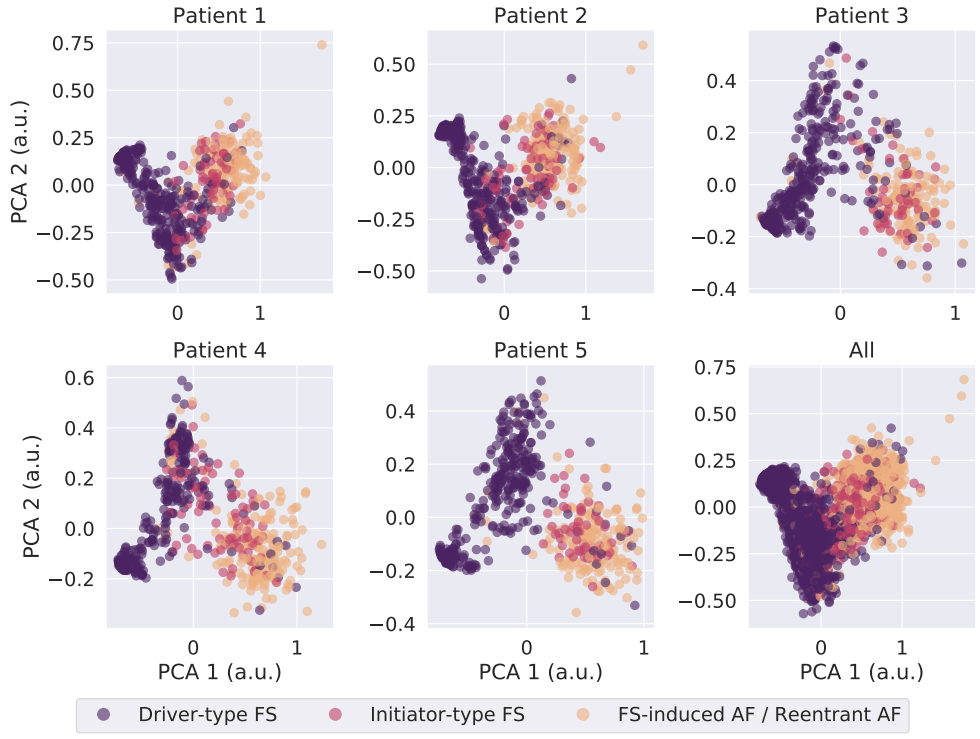
SO-BSS was able to capture the FS in most cases. This can be seen in Fig. 4.12 that the estimated dominant CL, extracted by SO-BSS with $K = 10$ from BSPMs, well approximated the ground-truth focal CL of the FS or its harmonics. The misalignment between the dominant CL and focal CL was mostly caused by reentrant activity during IFS. For some DFS with a CL of 150 ms, the dominant CL was estimated to be around 225 ms, which is the average of the focal CLs of 150 ms and 300 ms, the latter resulting from 2:1 conduction blocks. For 98.1% FS, their focal CL was estimated as source CL or its harmonic, with an absolute error ≤ 5 ms. Focal ablation increased the dominant CL of IFS (means of 196 ms (before) and 202 ms (after), p-value < 0.001, pairwise one-sided t-test).

4.3.5 Applying trained classifiers to predict patient's AF mechanisms

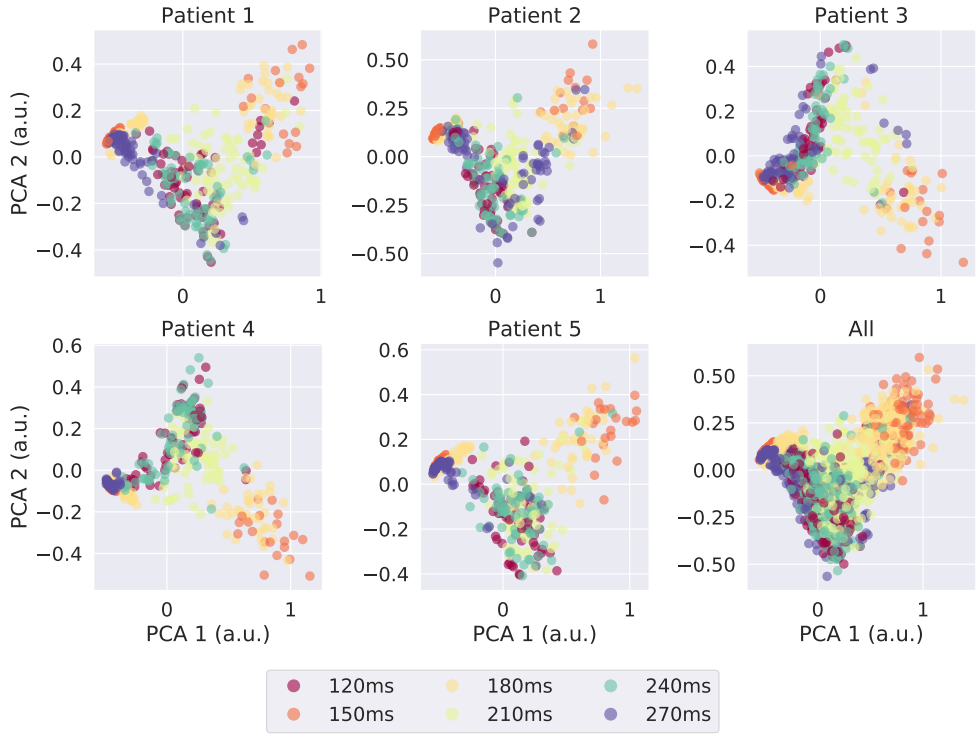
Having seen good classification results on our simulated data, the next step was to apply our trained classifiers on the clinical data set of 50 paroxysmal AF patients, in order to test if they bring in any predictive value for the outcome of ablation treatment. We used the classifiers trained on the BSPMs of the virtual cohort, with $K = 10$ for FS presence and AF sustainability and with $K = 1$ for FS location, to predict signal-level mechanisms from patient BSPMs. We then obtained the patient-level mechanisms by aggregating all signal-level mechanisms of the patient. The number of patients with each patient-level AF mechanism is shown in Table 4.5. All patients had an arrhythmogenic substrate. IFS were found in almost 80% of patients, and IFS from the LA or the PVs were found more often than from the RA.

The SO-BSS decomposition and the spatial pattern of an example patient test case are shown in Fig. 4.13 and Fig. 4.14. In Fig. 4.13, it can be seen that the first three extracted sources still had relatively high MaxAC which are above 0.85. The CL with the maximal ACF extracted from the V1-lead ECG signal, 188 ms, is similar to the CL with the maximal ACF in the first row, 191 ms, but the first source had a higher periodicity. The contribution of this first periodic source to the surface is also visible in Fig. 4.14, where the contribution from s_1 to the signal is also clearly seen. This serves as a quantitative validation of using our algorithm on patient signals.

The simplest AF mechanism was driven by an FS from a single atrium, which was also the most straightforward driver to map and ablate during a catheter ablation procedure. Therefore, we divided the patient cohort according to whether each patient contained an IFS originating from a single atrium.



(a)



(b)

Fig. 4.11. Two-dimensional PCA representation of MaxAC values. MaxAC values were obtained by SO-BSS with $K = 10$ over simulated BSPMs, on each patient and all pooled, for different colour-coded groupings: (a) all categories over all episodes, and (b) focal CLs over all FS episodes. Clustering based on categories and focal CLs can be seen for each patient. a.u.: arbitrary unit.

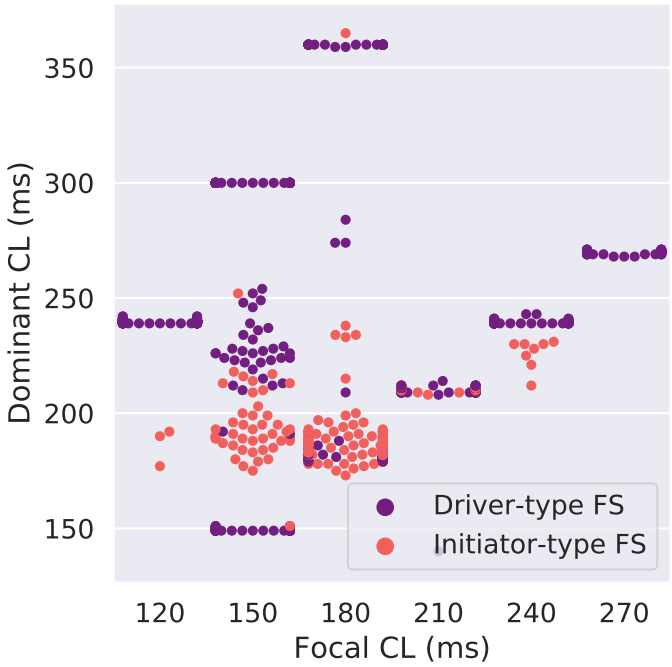


Fig. 4.12. A beeswarm plot of the estimated dominant CL and the ground-truth focal CL. Data points are coloured by FS type. CL was computed using $K = 10$ for SO-BSS from BSPM signals. Some samples were collapsed where possible for better visualization.

Table 4.5. Classification of patient data. Groupings and counts of patient-level mechanisms (by row), as well as the assignment of the patient group for survival analysis in Fig. 4.15, based on whether a patient contains IFS coming from a single atrium. This table omits the results of patients with DFS for clarity.

Grouping	IFS		Sustained AF without FS	Count
	from LA/PVs	from RA		
Group 1 (Single-atrial IFS)	✓		✓	18 (36%)
		✓	✓	9 (18%)
Group 2 (Other mechanisms)	✓	✓	✓	12 (24%)
			✓	11 (22%)
Count	30 (60%)	20 (40%)	50 (100%)	50 (100%)

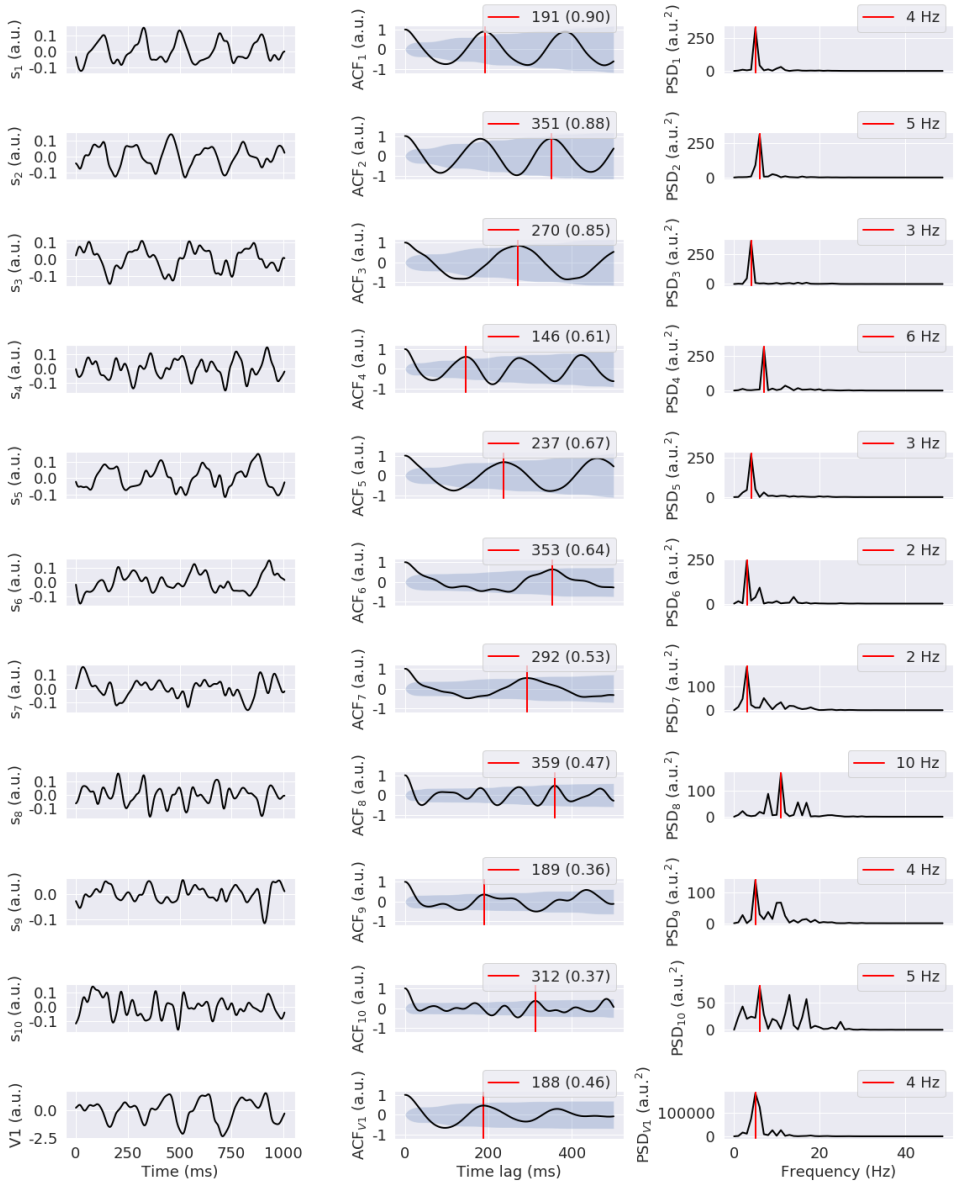


Fig. 4.13. V1-lead ECG signal and SO-BSS sources of estimated from a f-wave segment of a paroxysmal AF patient. The relatively high MaxAC of the first three sources showed that the periodic components were extractible from the patient signals. The legend is the same as Fig. 4.2.

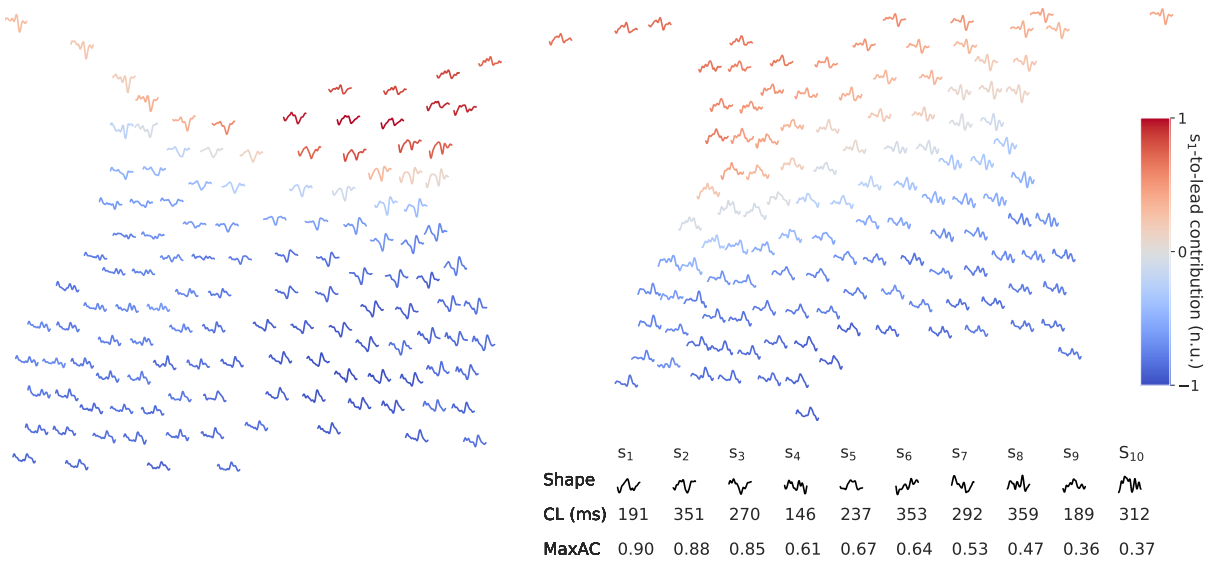


Fig. 4.14. Signals and estimated sources after channel-wise normalization of a BSPM signal of a paroxysmal AF patient, up to 300 ms. The legend is the same as Fig. 4.5. The similarity of the $\mathbf{s}_1$ with surface signals with a large contribution from $\mathbf{s}_1$ (in dark red) showed that it was reasonable to apply SO-BSS on BSPMs to extract common periodic source for patient signals.

To analyse how the AF recurrence likelihood can be predicted by covariates, the AF recurrence event was marked as a death event for survival likelihood regression. An entry in the data for survival analysis constitutes the time (number of months since the ablation) and event (whether the patient had experienced AF recurrence by that time). The patients were tracked until their last reported AF-free or the first reported AF recurrence event. Since AF recurrence might still happen after the last reported AF-free events but not reported, we only know the lower-bound of the AF recurrence, and the follow-up data is "right-censored", which is typical in clinical follow-up studies.

We calculated 36-month Kaplan–Meier curves as an estimation for AF-free likelihood on the right-censored data [202], and logrank tests on the difference of the two Kaplan–Meier curves, up to one-year, two-year and three-year follow-ups, using the lifelines Python package (version 0.26.0) [203], as shown in Fig 4.15. A higher post-ablation AF-free likelihood was observed for the patient subgroup with IFS originating from a single atrium than the other group, with p-values < 0.05 for all one-year, two-year and three-year follow-ups. This demonstrates that by classifying patient AF mechanisms using their pre-operative BSPMs, we successfully predicted a patient subgroup effectively targeted by the current standard procedure.

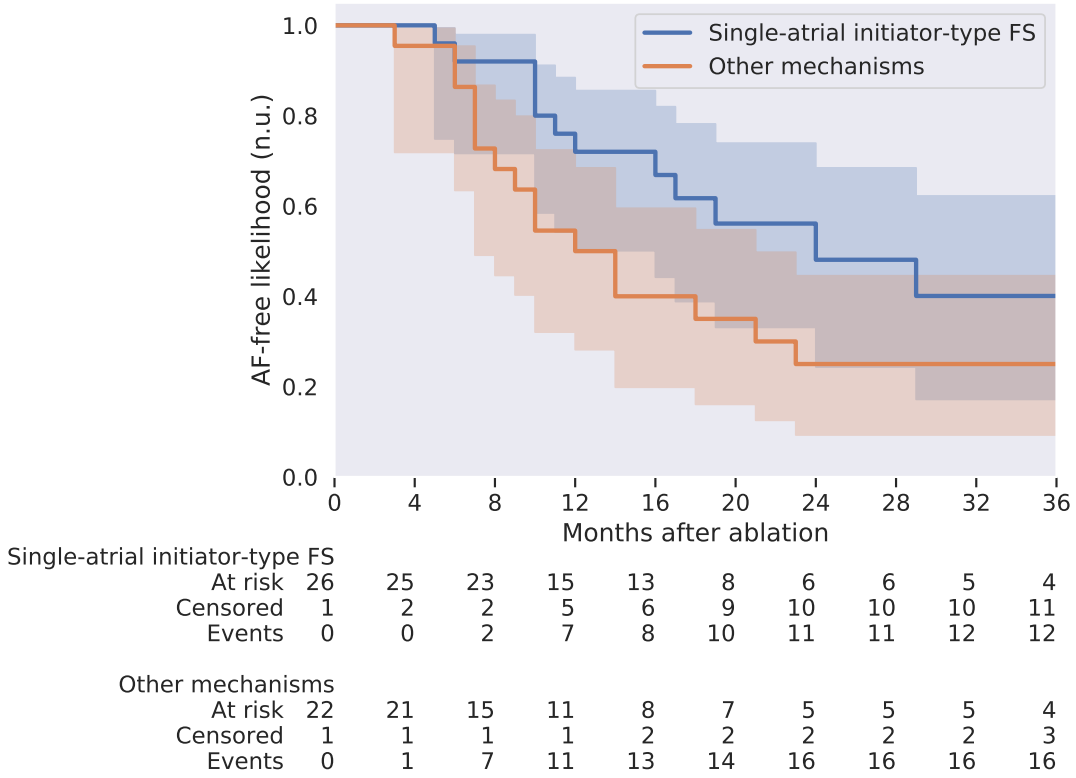


Fig. 4.15. Kaplan–Meier curves of two paroxysmal AF patient groups. Curves show the post-ablation AF-free likelihoods up to three years. The grouping was according to whether AF episodes in the patient were predicted as being driven by IFS originating from a single atrium. Shading shows the 95% confidence interval of the Kaplan–Meier curves. The group of single-atrial IFS had better one-year (p-value=0.017) , two-year (p-value=0.014) and three-year (p-value=0.025) AF-free outcomes than the other group, with p-values of logrank tests all < 0.05. n.u.: normalized unit. At a particular time shown vertically above, "At risk" and "Events" show the count of patients without and with AF recurrence up to that time, respectively, and "Censored" shows the number of patients who did not have AF recurrence up to that time, but no event was recorded afterwards.

4.4 Discussion

4.4.1 Mechanism-inspired classifiers

The results above suggested our simulations were reasonably representative of AF observed in patients, and thus, were ready to be used as the training data for non-invasive classifiers. They also showed that AF sustainability could be predicted from FS-driven episodes by two main features, focal CL and the presence of sustained rotors. These features are, however, usually extracted from invasively-acquired intra-cardiac electrograms. To translate these findings to non-invasive measurements, in order to detect the FS and the arrhythmogenic substrate, we approximated them with atrial sources extracted by SO-BSS from body surface potentials. As a result, the FS and the reentrant sources can be separated even when these two types of sources co-exist, such as in the case of IFS, which demonstrated usefulness in AF ablation targeting.

The focal CL was well approximated by the dominant CL, extracted by SO-BSS for the DFS and most of the IFS (Fig. 4.12). It was more advantageous to use the multi-lead SO-BSS than ACF or FFT analysis on the V1 lead, as SO-BSS extracts a spectrum of CLs. In addition, on one-second 1 kHz body surface signals, the resolution of FFT-extracted DF was 1 Hz, much lower than the resolution of 1 ms in our method. For the IFS which contained both FS and reentrant sources, their dominant CL increased after the removal of FS. This was consistent with the inverse relationship between the CL of AF waves and the number of AF drivers [114]. Such an observation strengthens faith in our model, and shows that our predictor could distinguish FS in spite of the presence of re-entrant sources.

The presence of sustained rotors was represented by MaxAC features. On a PCA-reduced space of all MaxAC features, a clear boundary can be seen between DFS and the other types, showing a distinctive difference between MaxAC features with and without an arrhythmogenic substrate, respectively (Fig 4.11(a)). The CL clusters in the MaxAC feature space (Fig 4.11(b)) also show that the MaxAC features encoded information about the focal CL, via their representation of atrial states, i.e. reentrant sources, FS, or both.

By using the contribution of each lead to the first periodic source ($\mathbf{s}_1$-to-lead contribution) during FS-driven episodes, we were able to predict whether the FS originated from the LA or the RA. We showed that this feature was robust to the changes of CLs, spatial heterogeneity, patient atrial meshes, but was sensitive of change of focal site, which made it a suitable feature to encode the spatial information. This feature was again inspired by the periodic nature of FS. An advantage of using BSPMs in inferring spatial information, compared to ECGs, was also proven by higher classification scores with BSPMs.

Our random forest classifier automatically combined the aforementioned heterogeneous features without additional re-scaling for classification. By doing so, a complex decision boundary between classification target classes was formed, whilst reducing overfitting by ensembling the results between multiple decision trees.

4.4.2 Robustness of classifiers

Our methods achieved high accuracy, precision and recall scores in Table 4.2, outperforming the AFFTr$_{2DF}$ [131] and NDI [15] features. We highlight that our non-invasive classifiers were robust to intra-patient variability, given a large variation of the atrial positions and shapes (Fig 3.4) and the changes in the vest positions and orientations (Fig. 4.10) In particular, the classification of FS presence and AF sustainability, as well as the estimation of the dominant CL, were robust to large changes to vest placement (up to 10 cm in translation and 20 degrees in orientation that were not represented in the training dataset. The heart-torso variability was not captured by the reconstructed periodic sources, and was "digested" by the transformation matrices during source reconstruction of SO-BSS. Neither the features of these two classifiers nor the dominant CL relied upon the transformation matrix. For other technologies however, such as ECGi, inferred atrial potentials from BSPMs could be quite sensitive to the relative positioning and orientation of the heart within the torso [204, 11]. As the joint diagonalization is robust to white noise, and there is no common structure of second-order statistics in it, our algorithm is resilient to recordings with white noise, as shown in our previous work [187].

It is worth noting that the ± 5 cm and ± 10 cm translation in the y-direction resulted in the hearts of five patients falling out of the boundary of the vest. This resulted in an inaccurate calculation of the BSPM using ϕ_e-recovery method. However, we still presented the result here to show the robustness of the method.

Our FS presence and AF sustainability classifiers were insensitive to the number of sources for SO-BSS when K was larger than a certain number, which was about 6 (Fig 4.9). The robustness of changing K means that prior knowledge about atrial sources is not required, and a smaller number of sources (< 6) can be selected for faster computation with moderate accuracy. To detect FS and AF sustainability, $K = 10$ sources are recommended with either ECGs or BSPMs, and $K = 1$ with BSPMs is recommended for predicting the focal site region. Furthermore, our method does not require imaging the heart or the torso, which accelerates the mapping process, and makes the tools accessible for both clinical and personal usage.

4.4.3 Related studies on AF inducibility from FS

Given the same pacing duration, only some FS with certain CLs and tissue conditions managed to induce self-sustained AF, suggesting that using the ectopic beat counts to predict for AF incidents in the previous works [160, 161, 162] could be improved by including a prediction of AF sustainability by our work.

The AFFTr_{2DF} feature was also developed on simulated episodes, but was only tested on FSs of 3 Hz and 5 Hz. This rarely induced complex patterns such as rotors or 2:1 atrial responses that were commonly observed in patients, confirmed by the morphology of FS in their results. In addition, their calculated power density spectrum, calculated from 500 ms of 500 Hz sampled signals, had a resolution of 0.01 Hz, much higher than the theoretical frequency resolution of FFT, which we were not able to reproduce. In comparison, our work not only considered a wide range of focal CL in FS, some of which co-existed with reentrant sources, but also added ACh regulation to further increase the spatial APD heterogeneity of the LA. Therefore, our classifiers, being trained and evaluated on more realistic simulations of FS and reentrant sources, were more robust.

The NDI feature, as an AF complexity metric already applied on AF patient signals, did not perform well on our tasks either, with a performance similar to the AFFTr_{2DF} feature. This may be due to the fact that their defined organized sources, captured by the three principal axes of PCA, could still include some organized reentrant sources.

Another simulation study [205] presented a classifier to detect PV versus non-PV drivers with the 12-lead ECG, which successfully predicted the acute success of PV isolation in 46 AF patients. Our AF sustainability classifier does not restrict the target procedure to PV isolation, and more importantly, detects an arrhythmogenic substrate that indicates AF susceptibility in the face of new AF triggers. This is more appropriate for predicting mid-term and long-term outcomes of a given treatment.

4.4.4 Limitations

The other type of IFS initiated reentry, anatomical macro-reentry, occurred much less frequently than DFS. The similar AF inducibility from an FS with a CL of 210 ms across all focal site regions, alongside a large inter-patient variability in the AF inducibility on this focal CL, suggests that global geometrical factors, shared by both atrial chambers, may need to be considered for predicting AF inducibility from FS. The follow-up records for paroxysmal AF patients contained missing entries and were retrospectively interpreted.

The ambiguity of the sign of the source is an intrinsic problem of the signal. In ACF

analysis, we do not need to care about the sign, but it may be vital to some applications, such as our second application, using the $\mathbf{s}_1$-to-lead contribution to estimate the focal site. We used an empirical method to refer to the sign of the maximal or mean value of the signal-to-source transformation matrix as the positive direction, but this unavoidably still adds ambiguity to the prediction of the focal site.

Chapter 5

Non-invasive atrial periodic source spectrum for AF condition representation

5.1 Introduction

In this chapter, we sought to extract features from pre-operative BSPM of 147 persistent AF patients, which can predict the ablation outcomes. Previously, we discussed the classification of FS and re-entrant sources using random forest classifiers with CL-MaxAC pairs, which were obtained from the body surface potentials with SO-BSS. The classifiers of "FS presence", "AF sustainability" and "FS location" were trained and tested using leave-one-out cross-validation on 2,977 simulated episodes. We showed that those classifiers were able to predict two cohorts with distinctive treatment effects on a paroxysmal patient dataset.

However, the mechanisms of persistent AF are more diverse and elusive, posing challenges in building realistic simulations that are representative of persistent AF patients. As the resultant machine learning models trained from the synthetic data heavily depend on whether the simulated signals are realistic, we explored machine learning models that could be trained directly on patient data, with an emphasis on ***simplicity*** and ***interpretability***. To cope with the limited number of patients in clinical studies, it is preferable to use a simple model with fewer parameters, so that the machine learning models can be generalizable to patient data. Simple models are also much easier to interpret in terms of physiological importance, and thus, it is easier to apply or adapt the methodology to data obtained from different sources. It is with these considerations that we replaced the random forest classifier model used in the last chapter with a linear classifier on a single variable, the simplest machine learning model, to predict the post-ablation outcomes for persistent AF patients.

The periodic content is electrophysiologically important, as it is associated with different phenomena such as FSs and rotors. For FSs and rotors, their frequency, or the periodic cycle length (CL), is likely drawn from a similar range, and the coexistence of several periodic components with different periodic CLs are informative to a mixture of different AF dynamics. In persistent AF patients, the highest dominant frequency (DF) content is spatially stable in corresponding atrial regions with varied frequencies [206, 207]. The non-invasive extraction of atrial periodic content motivated the development of an atrial periodic source spectrum (APSS), which is a function of maximal periodicity of the atrial sources over a prescribed range of CLs, and is more comprehensive than the highest DF. Moreover, APSS has a time resolution the same as the sampling frequency (such as 1 kHz in our case) regardless of the length of the signals. It is much higher than the 1 Hz precision achieved by the FFT on the f-wave segment of a single beat (typically one second or less) with the sampling frequency of 1 kHz. Similar to FFT, our metric of APSS can either be inferred from the BSPMs, or computed directly on intracardiac signals.

This chapter presents the development, validation and application of APSS extracted non-invasively from pre-operative ECG or BSPM, to enable machine learning to predict labels/events such as post-ablation AF recurrence, directly on patient data. We called APSS inferred non-invasively as ***non-invasive APSS***, and extracted directly as ***true APSS***. We show that the non-invasive APSS approximates the true APSS, which can be used to predict two subgroups with different long-term post-ablation AF recurrence amongst our dataset of persistent AF patients.

5.2 Methods

The estimation, validation and application of non-invasive APSS are shown in Fig. 5.1, and their details are described in the subsections as follows. In summary, the estimation of non-invasive APSS is obtained via CL-MaxAC pairs from a f-wave segment of body surface potentials, either BSPMs or ECGs. A patient-level APSS was obtained by aggregating all segment-level APSSs of a patient. The patient-level APSS was downsampled afterwards for the survival analysis on 147 persistent AF patients, going through ECGi-guided ablation as presented in Haïssaguerre et al. [41], where focal and reentrant drivers were first mapped by ECGi using the same (or very similar) set of surface recordings, and they were subsequently ablated until the slowing of local atrial activity. If AF was not terminated, linear ablation with LA roof and mitral isthmus was performed.

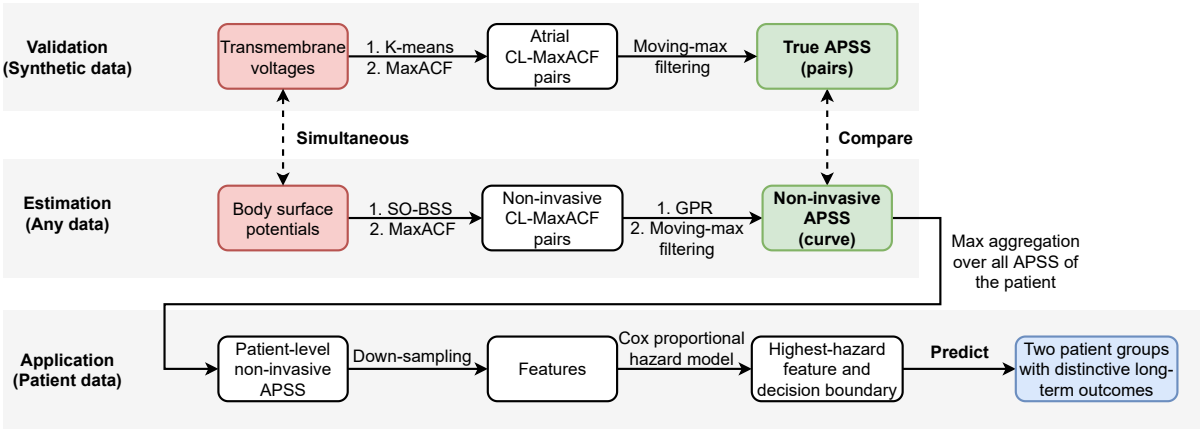


Fig. 5.1. Flowchart of the study of non-invasive APSS, the validation on a synthetic dataset, as well as an application of the APSS to predict two groups with distinctive long-term AF-free probability on a persistent AF dataset.

5.2.1 Estimation of non-invasive APSS

SO-BSS

First, as described in the previous chapter, SO-BSS was applied to extract a user-specified number of equivalent periodic sources from body surface potentials (either ECG or BSPM) to obtain CL-MaxAC pairs. As in the last chapter, we chose to extract $K = 10$ CL-MaxAC pairs to achieve a good balance between source estimation accuracy and time efficiency.

Estimation of APSS with a Gaussian process for regression (GPR)

In order to estimate MaxAC over a range of CLs, we used GPR over these K CL-MaxAC pairs to represent the MaxAC (target/output domain) over a continuous range of CL (input domain). The mean and the variance of the GPR prediction are taken as the estimation and uncertainty of the non-invasive APSS. This is based on the assumption that MaxAC is continuous in the CL domain, and the MaxAC decreasing on the query that CL is moving away from the true CL from a periodic source. A zero-mean GPR function was adopted as we assumed there is no periodic content (MaxAC equals to 0) on CLs far from all estimated CL-MaxAC pairs.

GPR [71] (section 2.3.1) was used to estimate an infinite unknown function $y_i = f(\mathbf{x}) + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma_n^2)$ with white noise represented by ϵ. The GP posterior

$$\mathbf{f}^*|\mathbf{X}^*, \mathbf{X}, \mathbf{f} \sim \mathcal{N}(m(\mathbf{f}^*), cov(\mathbf{f}^*)) \tag{5.1}$$

were computed to represent the APSS from observations, $(\mathbf{X}, \mathbf{y}) = \{(\mathbf{x}_i, y_i)|i = 1, 2, \ldots, K\}$, which were the K CL-MaxAC pairs obtained from SO-BSS. $\mathbf{X}$ is an $K \times 1$ training input matrix, and $\mathbf{y}$ is a vector of all scalar outputs. $\mathbf{X}^*$ denotes the $|T| \times 1$ test input matrix, where we used $T = \{100, 101, \ldots, 400\}$ms. The mean $m(\mathbf{f}^*)$ and variance $cov(\mathbf{f}^*)$ of the unknown function, the APSS, can be computed using the following formulae (see also Eq. (2.18) and (2.19)):

$$m(\mathbf{f}^*) = \kappa(\mathbf{X}^*, \mathbf{X})(\kappa(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I)^{-1}\mathbf{y} \tag{5.2}$$

$$cov(\mathbf{f}^*) = \kappa(\mathbf{X}^*, \mathbf{X}^*) - \kappa(\mathbf{X}^*, \mathbf{X})(\kappa(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I)^{-1}\kappa(\mathbf{X}, \mathbf{X}^*) \tag{5.3}$$

The covariance function $\kappa(\cdot, \cdot)$ is crucial for GPR estimation. We chose a Matérn kernel [71] with $\nu = 3/2$ for the covariance function, a commonly used isotropic kernel that is a product of an exponential and a polynomial of one degree. The kernel value $\kappa(\cdot, \cdot)$ only depends on the distance between input pairs $\kappa(\mathbf{x}, \mathbf{x}') = \kappa'(r = |\mathbf{x} - \mathbf{x}'|)$, and has the form of

$$\kappa'(r; l) = (1 + \frac{\sqrt{3}r}{l}) \exp(-\frac{\sqrt{3}r}{l}) \tag{5.4}$$

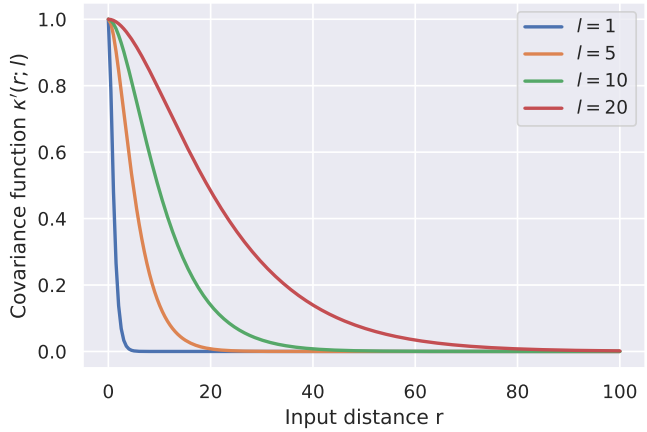
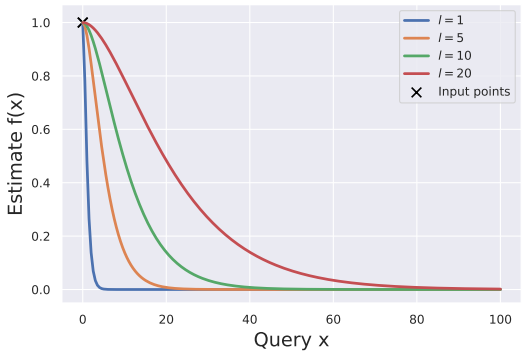


Fig. 5.2. The kernel value $\kappa'(r; l)$ based on input distance r, with different hyperparameter l. A larger l entails slower descend on the covariance function.

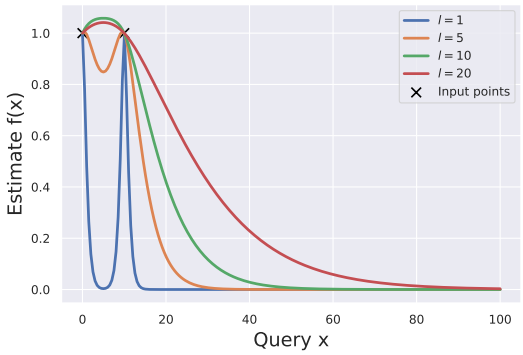
A visualization of the kernel is seen in Fig. 5.2, which shows that the larger l is, the slower the kernel value descends when the input distance r increases, and therefore, the smoother the GPR estimation would be. This can be further seen from Fig. 5.3, with the estimation of the GPR mean (Eq. (5.2)) for an isolated input point, as well as paired points of different distances. For an isolated input point in Fig. 5.3(a), the curve resembles that of the covariance function $\kappa'(r)$ in Fig. 5.2. For $r = 5$ in Fig. 5.3 (b), for $l = 10$ and 20, there is an arc connecting the two input points, but for smaller $l = 1$ and 5, the two points are connected with a declining curve. Increasing the distance

r between the two inputs, a larger l is needed to have an upward curve connecting them (Fig. 5.3 (b-d)).

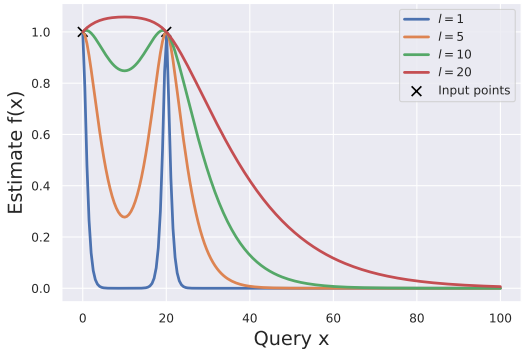
As all hyperparameters of the covariance function are updated during the optimization process of maximizing the marginal likelihood $\log p(\mathbf{y}|\mathbf{X}, \theta)$ (Eq. (2.21)), a bound of the characteristic length-scale l which was set to reflect our belief for modeling the true APSS. We assumed periodic sources with similar CLs should be considered as having the same periodic content, and, for CLs with high MaxAC, the MaxAC diminishes as moving away from these CLs. We then set the length scale l to be bounded by $1 \leq l \leq L$ with $L = 5, 10, 20$ and 30 as candidates. Further consideration of choosing the boundary is presented in the result section (5.3.2 and 5.3.3).



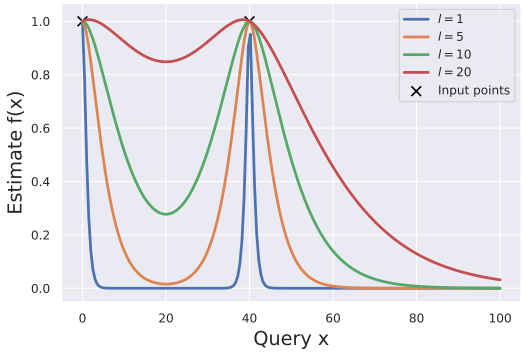
(a) An isolated input point.



(b) Input pairs with distance $r = 10$.



(c) Input pairs with distance $r = 20$.



(d) Input pairs with distance $r = 40$.

Fig. 5.3. Estimations of the GPR using different hyperparameter l. The distance $r = 10, 20$, and 40 for (b-f), and the distance $l = 1, 5, 10$ and 20 are typical points of the specified bound for l.

In addition, the standard deviation σ_n of the white noise $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ as specified in Eq. (5.2) and (5.3) was set to be bounded by $0.001 \leq \sigma_n \leq 0.05$ in estimating MaxAC by the SO-BSS step.

We used the scikit-learn [200] Python package (version 0.22.1) to implement the GPR.

Moving-max filtering on CL-MaxAC pairs

We observed that among the CL-MaxAC pairs from a f-wave, there are still entries with similar CLs but varied degrees of MaxAC (see the left panels of Fig. 5.4). Therefore, before applying GPR with CL-MaxAC pairs as input, we used a moving max filter to keep only the pairs with MaxAC being the maximum over the adjacent CL within a radius of 10 ms. A demonstration of this step and the difference between before and after this filtering step are shown in Fig. 5.4.

5.2.2 Validation of non-invasive APSS

Mean shapes

In order to demonstrate that non-invasive APSS estimated by GPR is able to represent the true APSS, we calculated the distribution of MaxAC over CL from the atria for comparison. For a time series of transmembrane voltage on an atrial mesh of N nodes, the computation of ACF with time lags of $1, 2, \dots T$ ms has a complexity of

$$O(N \times (1^2 + 2^2 + \cdots + T^2)) = O(N \times \frac{T(T+1)(2 \times T + 1)}{6}) = O(N \times T^3)$$

which is too expensive to evaluate in its exact form. Therefore, we measured the function of MaxAC over CL by first clustering transmembrane voltages of all atrial nodes with similar morphologies using a minibatch K-means algorithm [208] to perform K-means clustering [209] on a large scale. K-means is an unsupervised learning technique to aggregate data points, with similar patterns by iteratively updating the means of each cluster of the data points.

True APSS

We then evaluated CL-MaxAC pairs of the mean morphology of each cluster. This reduced N in the complexity term to a much smaller M. We set $M = 10$ for extracting CL-MaxAC pairs from the LA and from the RA, respectively. The $M = 10$ was chosen to be the same as the number of non-invasive CL-MaxAC pairs, which provided a similar degree of averaging effect to the transmembrane voltages in each atrium. In addition, this choice helped save computational time. We then concatenated these 20 pairs to constitute the true CL-MaxAC pairs of the atria, ordered by their CL. Similar to how we dealt with the CL-MaxAC pairs extracted from the BSPM, we applied a moving max filter to keep the maximum of CL-MaxAC pairs at a vicinity of 10 ms (Fig. 5.4). We denoted the final CL-MaxAC as $\mathbf{y}_{atria}$, which are datapoints sampled from the true APSS.

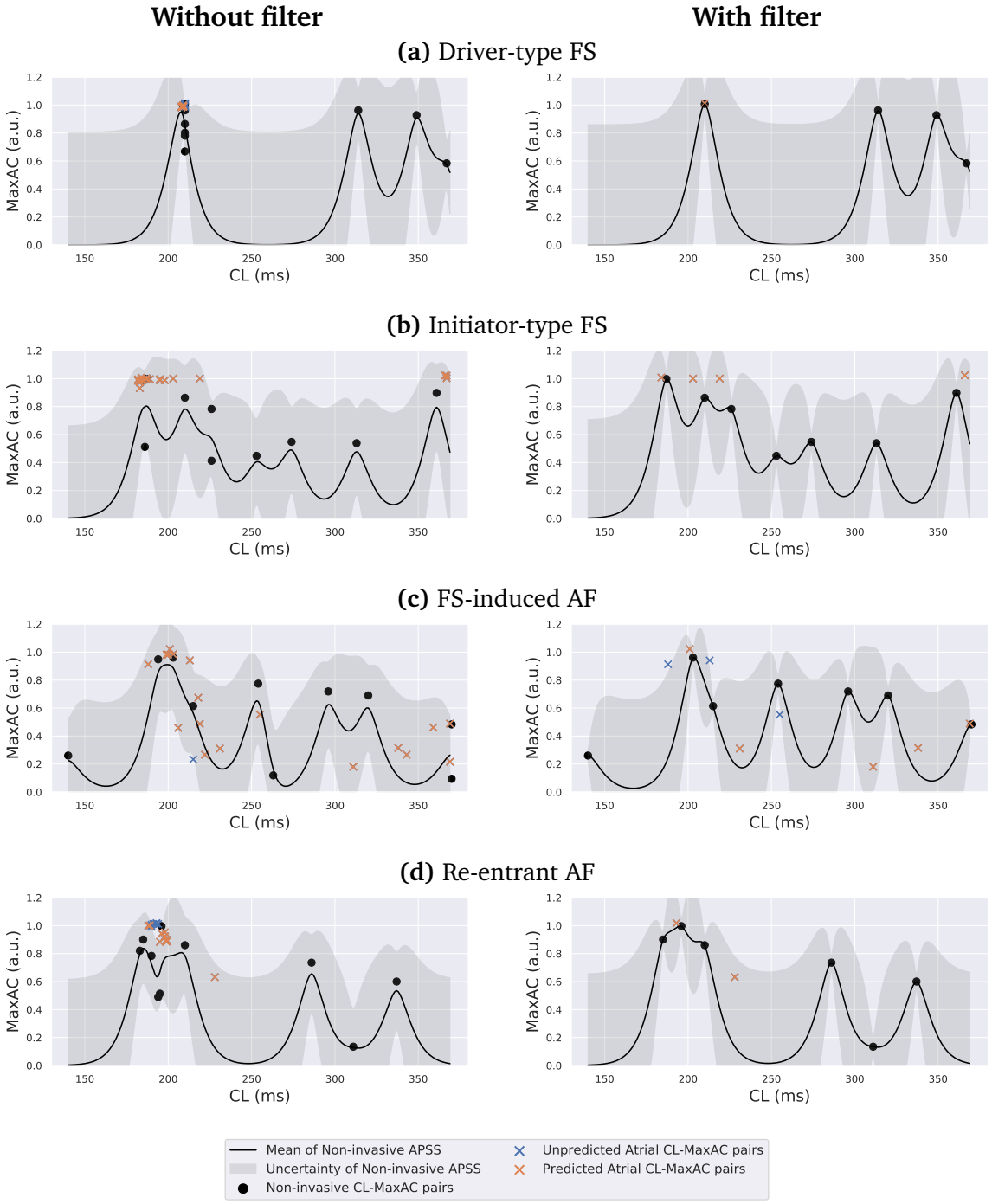


Fig. 5.4. Examples of non-invasive APSS (black curve) with uncertainty bound (gray shaded area) as prediction for the atrial APSS (crosses), before and after applying filters to both non-invasive CL-MaxAC pairs (dots) and the atrial CL-MaxAC pairs (crosses), within $1 \leq l \leq 10$, on (a) driver-type FS, (b) initiator-type FS, (c) FS-induced AF, and (d) re-entrant AF. The un-predicted and predicted truth pairs refer to those that are included or not predicted by APSS reconstructed from the BSPM (Eq. (5.12)). The gray shaded area shows the uncertainty boundary of APSS in Eq. (5.3). a.u.: arbitrary unit.

5.2.3 Application of non-invasive APSS

Aggregation for patient-level non-invasive APSS

To represent the state of a patient, we utilized a patient-level APSS, obtained by aggregating the maximal MaxAC for each CL interval across all APSSs of a patient, as shown in Fig. 5.5.

Downsampling patient-level non-invasive APSS

For survival regression, we down-sampled the APSS by taking the maximum of the MaxAC for every 10 ms of CL, as features (covariates). For CL ranges from 100 to 400 ms that we adopted in the study, this amounts to 30 features in total. Each feature represents the maximal MaxAC at an interval, and we named each feature by the smallest CL of the interval they represented. This process is illustrated in Fig. 5.5.

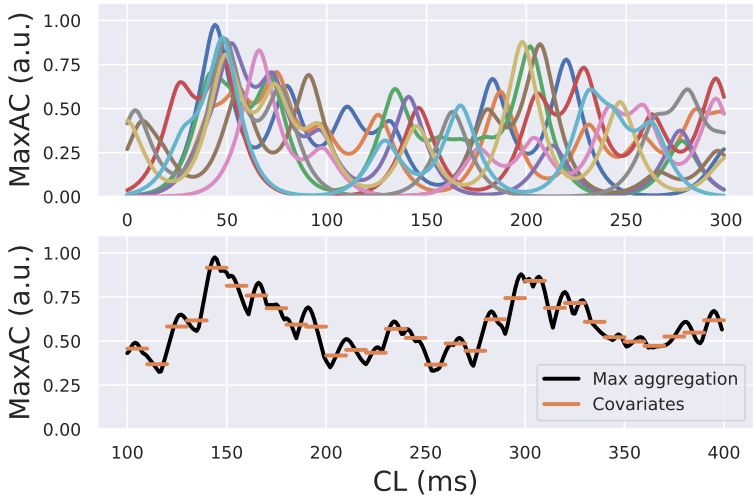


Fig. 5.5. Method of building the patient-level non-invasive APSS after applying GPR by aggregation across all non-invasive APSS of the patient, and downsampling to build features. The top plot shows the original non-invasive APSS estimation (lines), and the bottom plot shows the max aggregation of non-invasive APSS (black), and the covariates by mean MaxAC of every 10 ms (orange). a.u.: arbitrary unit.

Survival likelihood regression

The same as the last chapter (Section 4.3.5), the AF follow-up data is right-censored. The survival likelihood function $S(t) = P(T > t)$ is a non-increasing function over time variable t which defines the probability of survival on a random time $T > t$. The hazard

function $h(t)$ at time t denotes the probability of the death event occurring at time t, which has the following relationship with the survival likelihood function,

$$h(t) = -\frac{S'(t)}{S(t)} \tag{5.5}$$

where $S'(t)$ is the derivative.

We used the downsampled patient-level non-invasive APSS features as covariates for survival function regression. The n covariates were represented as $\{\mathbf{z}_i | i = 1, 2, \ldots, n\}$, where each $\mathbf{z}_i$ is a vector aggregating the scalar values of all patients for covariate i, and z_{ij} shows the value of this covariate i for patient j. We used an $n \times N_{Patient}$ matrix $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n]^T$ to show the values of all covariates, where $N_{Patient}$ is the number of patients. We also used $\mathbf{Z}_j$ to denote all n covariates of patient j.

Some of the most common survival regression models, such as the Cox's proportional hazard model [210], represent the hazard rate $h(t|\mathbf{Z})$ as a function of t and some covariate variable $\mathbf{Z}$. The Cox's model assumes the hazard ratios of all n covariates $\mathbf{Z}$ are time-invariant, and represents the total hazard rate as

$$h(t|\mathbf{Z}) = b_o(t) \cdot \exp(\sum_{i=1}^{n} b_i(\mathbf{z}_i - \bar{\mathbf{z}}_i)) \tag{5.6}$$

where b_0 is the baseline hazard function that is the same for all covariates, the partial hazard function b_i denotes the hazard function for a covariate i, and $\exp$ is the exponential function.

In survival analysis, the event of AF recurrence for a patient j at a time t is defined as $t = X_j$. The patient set $R(t) = \{j | X_j \geq t\}$ is called the risk set at time t, which denotes the set of patients "at risk" of AF recurrence at that time. The probability of someone (such as patient j) from the risk set $R(t = X_j)$ reporting AF recurrence at a particular time X_j is represented as a conditional probability as

$$\mathcal{LL}_j(\beta, \mathbf{Z}) = \frac{h(X_j|\mathbf{Z}_j)}{\sum_{l \in R(X_j)} h(X_j|\mathbf{Z}_l)} \tag{5.7}$$

where β denotes the hazard functions of all covariates.

Using the Cox proportional hazard model in Eq. (5.6) to the above equation, we obtained the log partial likelihood $\mathcal{LL}(\beta, \mathbf{Z})$ as a product of $\mathcal{LL}_j(\beta, \mathbf{Z})$ for all N_r patients with observed AF recurrence events, which is

$$\mathcal{LL}(\beta, \mathbf{Z}) = \Pi_{j=1}^{N_r} \mathcal{LL}_j(\beta, \mathbf{Z}) \tag{5.8}$$

$$= \Pi_{j=1}^{N_r} \frac{b_0(X_j) \exp(\sum_{i=1}^{n} b_i(z_{ij} - \bar{z}_{ij}))}{\sum_{l \in R(X_j)} b_0(X_j) \exp(\sum_{i=1}^{n} b_i(z_{il} - \bar{z}_{il}))} \tag{5.9}$$

$$= \Pi_{j=1}^{N_r} \frac{\exp(\sum_{i=1}^{n} b_i(z_{ij} - \bar{z}_{ij}))}{\sum_{l \in R(X_j)} \exp(\sum_{i=1}^{n} b_i(z_{il} - \bar{z}_{il}))} \tag{5.10}$$

where $\bar{z}_{il}$ is the estimated value of z_{il}, and the baseline hazard ratio $b_0(X_j)$ is eliminated in Eq. (5.10).

We also added an L1 penalty term with a coefficient of $\lambda = 0.01$, in order to shrink coefficients of less importance to the regression target. This gives the final form of the optimization target $\mathcal{LL}^*(\beta, \mathbf{Z})$ for all n covariates.

$$\mathcal{LL}^*(\beta, \mathbf{Z}) = \mathcal{LL}(\beta, \mathbf{Z}) - \lambda ||\beta||_1 \tag{5.11}$$

Survival estimation using the selected feature For survival likelihood estimation [202], when it is estimated using multiple features, cross-validation has proved to reduce bias in accuracy estimation [211]. Therefore, we used 10-fold cross-validation to train and test the performance. In each fold, about 90% of patients were used to select the most relevant feature with the largest hazard ratio, as well as a cut-off MaxAC threshold θ from the training dataset. The feature and the threshold was used to divide the test patients into the lower (MaxAC $\leq \theta$) and higher (MaxAC $> \theta$) groups. The lower and higher groups of all folds were then concatenated to form the total lower and higher groups, and Kaplan-Meier curves [202] were plotted for the final total lower and higher groups.

We used the lifelines Python toolbox (version 0.26.0) [203, 212] for the optimization and plotting of the Cox proportional hazard model and the Kaplan-Meier curves.

5.3 Results

5.3.1 Comparison between with and without SO-BSS

As periodic content is usually extracted by FFT, we demonstrated the role of SO-BSS in our dataset. We evaluated the non-invasive APSS extracted from BSPM with SO-BSS and GPR, and the following two candidates: (1) ***MaxAC of BSPM***, which directly computes MaxAC of CL on each lead of the BSPM, and aggregated the highest MaxAC of each CL across all channels; and (2) ***MaxFFT of BSPM***, using the amplitude of the FFT over CL (=1000 / frequency, unit: ms), and aggregated the maximum computed across all channels. The ACF of the BSPM underestimated the MaxAC at CLs around 200 ms (Fig. 5.6(a)), as some periodic sources were blurred when projected to the torso. The more irregular and complex the atrial electrical waves are, the more that SO-BSS is needed in revealing the periodic sources. FFT (red dash line) missed the spike of MaxAC at the CL of 300 ms (Fig. 5.6(b)), due to a limited precision using FFT

for one-second signal at a sampling frequency of 1 kHz. However, SO-BSS extracted the periodic structure among multi-channel signals.

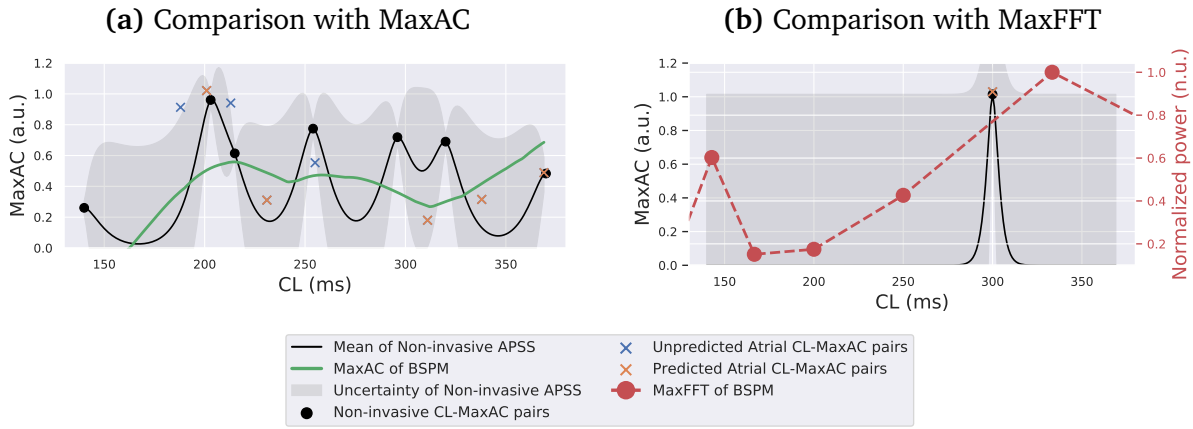
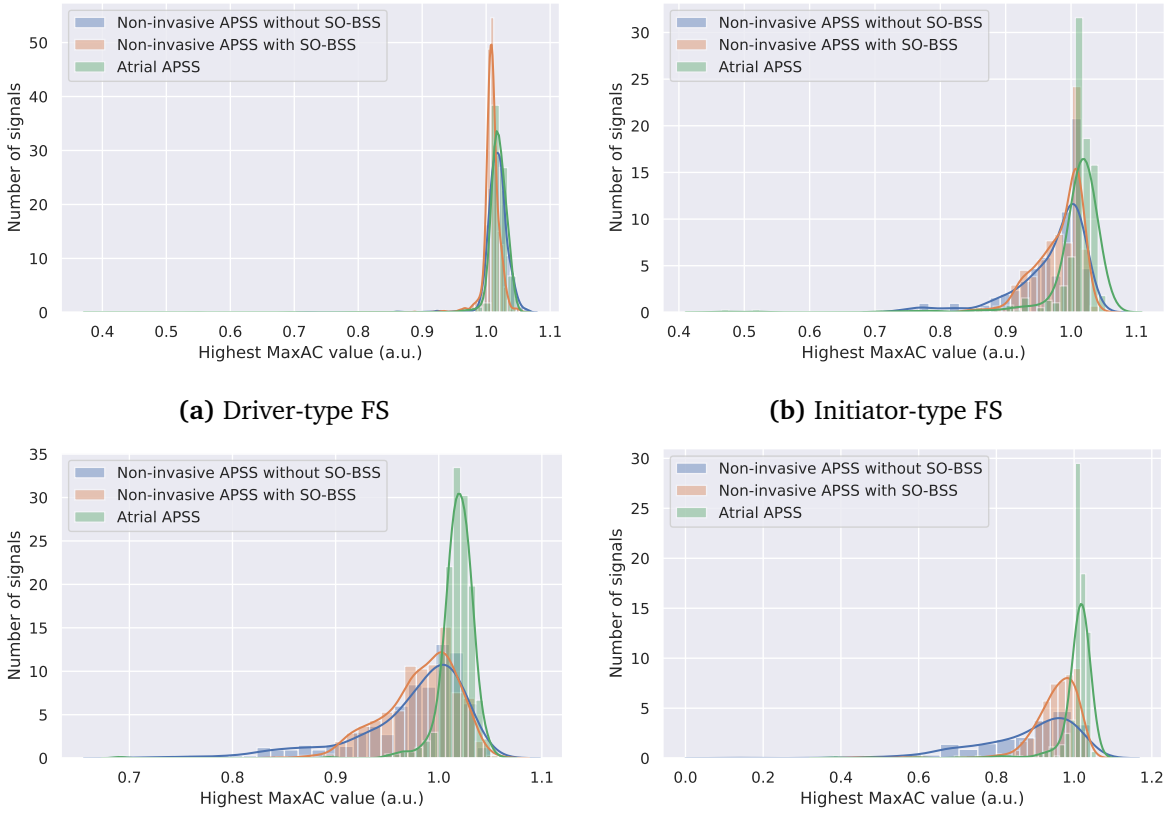


Fig. 5.6. Comparison between non-invasive APSS extracted from BSPM with SO-BSS and GPR (black curve), versus directly from BSPM without SO-BSS, using (a) MaxAC (green curve) or (b) MaxFFT (red dash curve). The SO-BSS points and the truth were both filtered by the moving max filter. The prediction shows the non-invasive APSS estimated with a hyperparameter bound $1 \leq l \leq 10$. MaxAC and MaxFFT of BSPM represents the maximum of the MaxAC value, or the normalized FFT amplitude over each channel of BSPM. (a) shows the comparison on a simulated signal of FS-induced AF with a pacing CL of 150 ms, where the ACF of the BSPM (green curve) underestimated the MaxAC at CLs around 200 ms. (b) shows the comparison on a simulated signal of driver-type FS, where FFT (marking CLs calculated from 1000 / frequency, unit: ms) missed the spike of MaxAC at the CL of 300 ms. a.u.: arbitrary unit. n.u.: normalized unit.

The distributions of the highest MaxAC values for the three methods in Fig. 5.7 further demonstrate the necessity of SO-BSS in extracting periodic sources from the body surface potentials. For complex cases (those except for driver-type FS), the SO-BSS recalled the highly periodic components better. The significant improvement in extracting periodic sources is the most prominent for FS-induced AF, which had the most mixed local and global wave dynamics, as each episode recording the shift from FS-dominant to self-perpetuated. As the dispersion between the two distributions in persistent AF datasets in Fig. 5.8 is similar to the degree of dispersion in Fig. 5.7 (d), the SO-BSS is expected to better extract the periodic content of the persistent AF dataset.

5.3.2 Comparison between non-invasive and true APSS

We used $\{(\mathbf{x}_i^*, y_i^*)\} = \{(\mathrm{CL}_i^{atria}, \mathrm{MaxAC}_i^{atria}) | i = 1, 2, \dots\}$ to denote the true APSS pairs extracted from the atria derived from transmembrane voltages. To compare the non-invasive APSS with the true CL-MaxAC pairs, we calculated the overlap between the uncertainty bounds (± 1 standard deviation) of the non-invasive APSS $m(\mathbf{x})$ and the true



(a) Driver-type FS

(b) Initiator-type FS

(c) Re-entrant AF

(d) FS-induced AF

Fig. 5.7. The highest MaxAC value of all sources estimated from BSPM with SO-BSS and GPR (orange), and extracted directly from BSPM without SO-BSS (blue), and those computed directly from transmembrane potentials (green) for different simulation categories (a-d) of the synthetic dataset. a.u.: arbitrary unit.

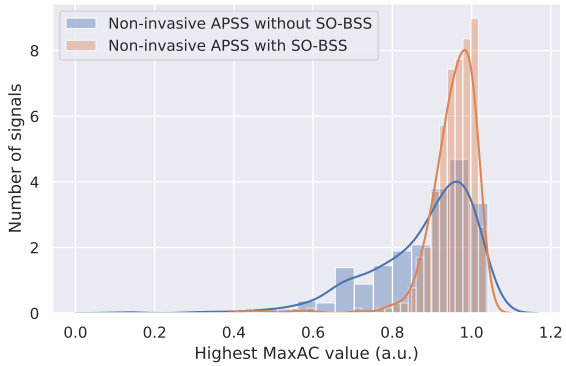


Fig. 5.8. The highest MaxAC estimated with SO-BSS and GPR (orange), and extracted directly without SO-BSS (blue) from BSPM of persistent AF patients. a.u.: arbitrary unit.

APSS. This is denoted by

$$Coverage = \frac{|\{(\mathbf{x}_j^*, y_j^*) : ||y_j^* - m(\mathbf{x}_j^*)|| \leq cov(\mathbf{x}_j^*)\}|}{|\{(\mathbf{x}_i^*, y_i^*)\}|} \tag{5.12}$$

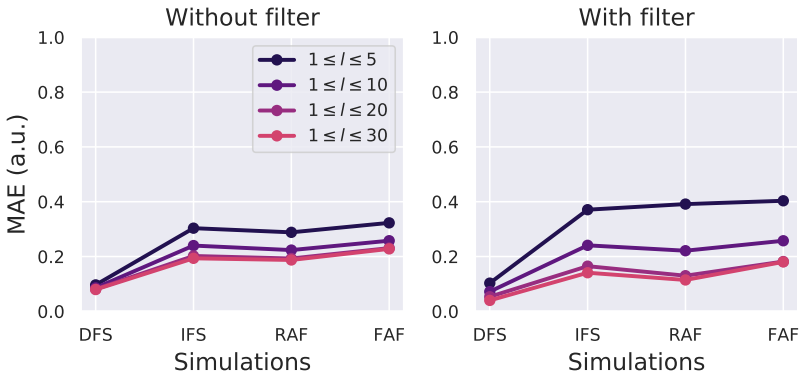
where $||\cdot||$ denotes the absolute value and the $|\cdot|$ denotes the cardinality. We also evaluated the prediction accuracy by the mean absolute error (MAE) between the estimated APSS and the true APSS at $\mathbf{x}^*$,

$$MAE = \sum_i ||y_i^* - m(\mathbf{x}_i^*)|| \tag{5.13}$$

The visualization of comparison between non-invasive APSS and the true atrial periodic sources are shown in Fig. 5.4. Non-invasive APSS curves are represented by atrial periodic sources after applying the moving max filters. For example, in Fig. 5.4 (d) before filtering, there were a few CL-MaxAC data points that are vertically scattered around CL of 195 ms, causing the prediction curve to be dipping at that location. However, this was corrected by a moving max filter, which preserved only the CL-MaxAC datapoint with the highest MaxAC, resulting in a peak at 195 ms.

The moving max filter improved both MAE and coverage significantly, as show in the Fig 5.9. The APSS for DFS simulations were all precisely predicted by the non-invasive APSS, with a MAE of 0.05 and the coverage of GPR almost 100%. For simulations of other AF mechanism categories, after filtering, the prediction of APSS achieved a fair accuracy with MAE below 0.2. The coverage also increased from around 60% to around 80%, with $1 \leq l \leq 10$ receiving the maximal coverage. Although the magnitude of the MaxAC was still somewhat compromised, our non-invasive APSS is a good approximation for the true APSS.

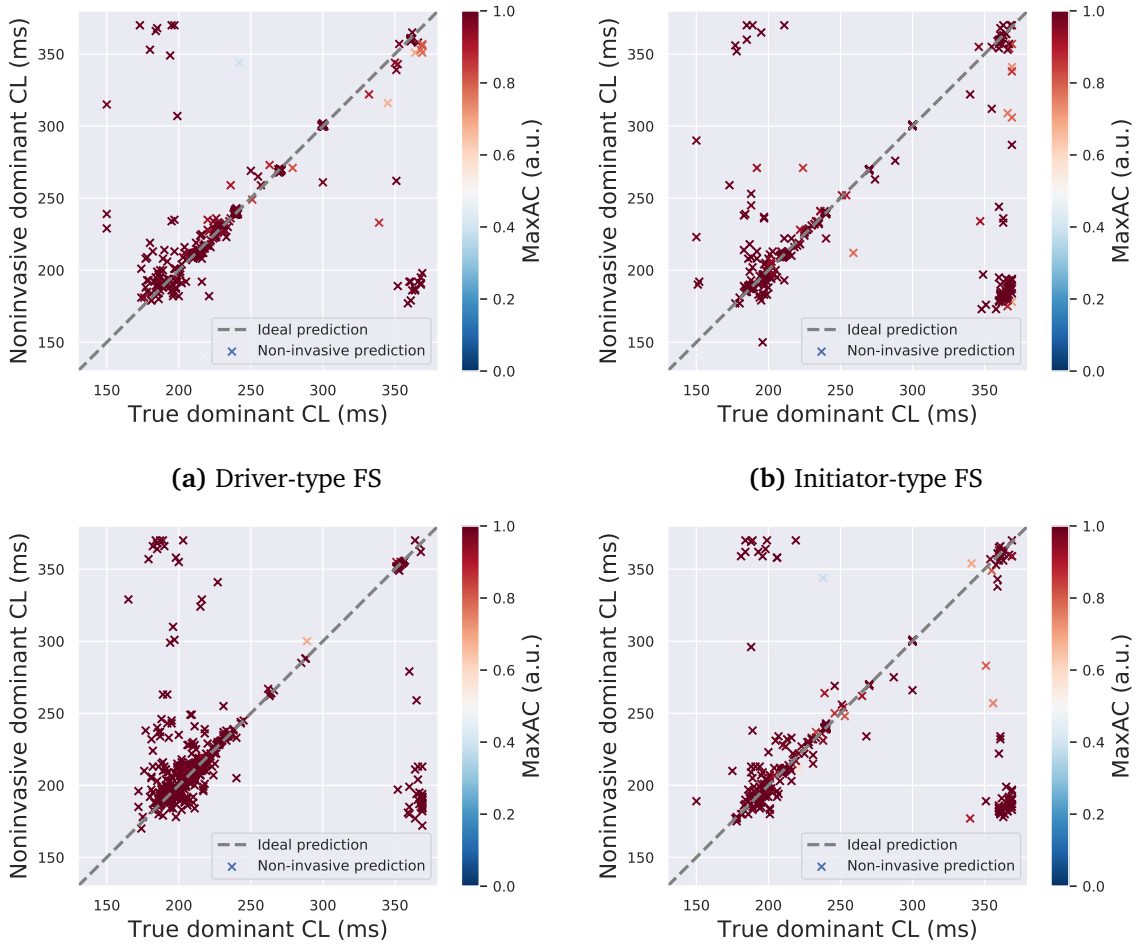
We also compared the dominant CL extracted from the non-invasive CL-MaxAC pairs and true CL-MaxAC pairs using the CL receiving the maximal value of MaxAC. The comparison is shown in Fig. 5.10. Compared to the grey dashed curve showing the ideal CL estimation, there was little systematic bias between the dominant CL estimated from the torso, which was a good advantage over DF analysis. The largest error was due to that some CLs were estimated as their harmonious components, as they also had a similarly large MaxAC amplitude. Therefore, we also evaluated dominant CL with estimation and ground truth both $\leq$ 300 ms, to exclude the effect of the harmonious components. All categories had similarly small estimation errors, with the total mean absolute error as 3.5 $\pm$ 9.7 ms.



(a) Mean absolute error

(b) Coverage

Fig. 5.9. The evaluation of (a) mean absolute error (MAE) and (b) coverage of ground truth (Eq. (5.12)**) of non-invasive APSS extracted from BSPM with SO-BSS and GPR for estimation of true APSS, on the synthetic dataset before and after applying a moving-max filter on both prediction and truth.** Filtering improves the estimation by reducing error and improving coverage. (a) The more relaxed the upper boundary is, the smaller the error is. (b) non-invasive APSS estimated with $1 \leq l \leq 10$ received the maximal coverage. DFS: driver-type FS. IFS: initiator-type FS. FAF: FS-induced AF. RAF: re-entrant AF. a.u.: arbitrary unit. n.u.: normalized unit.



(a) Driver-type FS

(b) Initiator-type FS

(c) Re-entrant AF

(d) FS-induced AF

Fig. 5.10. Comparison of the dominant CLs extracted from non-invasive CL-MaxAC pairs and atrial CL-MaxAC pairs, over four categories. The total mean absolute error of the CL is 13.8 ± 42.1 ms. When evaluated only for estimated and ground-truth dominant CLs both $\leq$ 300 ms, the total mean absolute error is 3.5 ± 9.7 ms. The mean absolute error for estimating the CL for each categories are (a) DFS: 3.6 ± 20.1 ms. (b) IFS: 38.0 ± 66.9 ms, (c) RAF: 20.2 ± 45.6 ms, and (d) FAF: 35.0 ± 65.0 ms. We also evaluated dominant CL with estimation and ground truth both $\leq$ 300 ms, the mean absolute error are (a) DFS: 1.4 ± 6.1 ms (b) IFS: 6.8 ± 16.5 ms, (c) RAF: 7.3 ± 11.5 ms, and (d) FAF: 4.8 ± 10.1 ms.

5.3.3 Additional considerations in selecting the GPR kernel hyperparameters

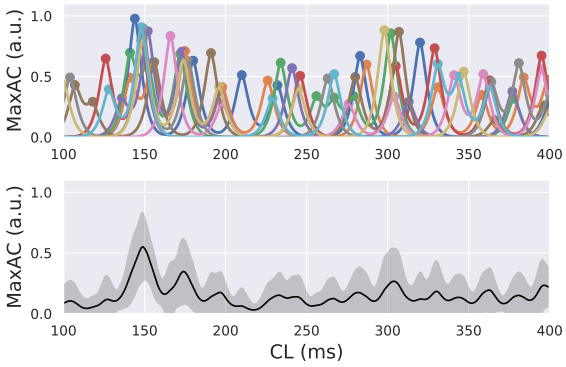
We present some additional consideration on selecting the GPR kernel hyperparameters, by cross-referencing APSS extracted from multiple f-waves of the same patient. As shown in Fig. 5.11, patient-specific patterns of APSS can be extracted through averaging APSS from different signals, where the left column shows a patient with a peak between CLs of 100 – 200 ms, and the platform-shaped pattern between CLs of 220 – 270 ms, resulted from at least two periodic sources with similar CLs, are shown in the second patient in the right column. The consistent intra-patient patterns from different AF episodes of the same patient suggest the presence of underlying patient-level features.

For AF patients with more complex AF episodes, where AF dynamics varied between different recordings, we used K-means clustering [209] using the Euclidean distance as distance metrics. Here, the number of clusters was selected by the silhouette analysis [213], commonly used for evaluating the assignment of clustering labels by how close a data point is to the other data points of the same cluster, compared to the data points of different clusters. An example of such way to find several patterns from a patient is demonstrated in Fig. 5.12, where clustering into 6 clusters had the highest silhouette score in Fig. 5.12(a). The APSS curves of different clusters (b-g) demonstrate the high similarity with other curves within the same cluster. The curves in Fig. 5.12(c)(e) exhibit peaks of different CLs, and are distinctive with other clusters.

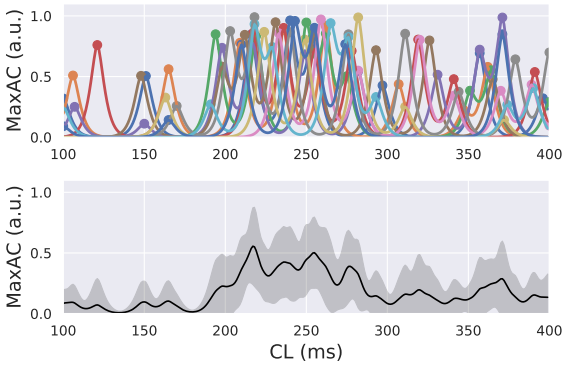
The effect of choosing different upper bounds for the time-scale l is also visible via plots of pair-wise correlation between the downsampled patient-level APSS features in Fig. 5.13. With a higher upper bound for l, the correlation plot has a less distinctive boundary with adjacent features. As this upper bound of l goes up, the correlation plot is increasingly closer to the correlation without SO-BSS in Fig. 5.13 (d). However, a negative correlation between 130 – 150 ms with 200 – 230 ms are more prominent with a higher upper bound. Based on evaluation on the simulation set, as well as the consideration to reveal local and common intra-patient patterns and feature correlations, we thereby have chosen $1 \leq l \leq 10$ as the hyperparameter bound.

5.3.4 Post-ablation AF recurrence prediction using APSS

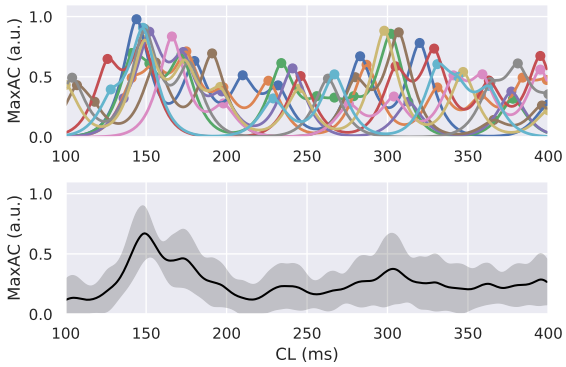
We used features derived from downsampled patient-level APSS to present the AF condition of each patient. Each feature is the highest MaxAC value corresponding to a CL interval of 10 ms, and were used as features for the Cox proportional hazard model, which predicts the hazard ratios associated with each feature, as shown in Fig. 5.14 for non-invasive APSS estimated with SO-BSS and GPR ($1 \leq l \leq 10$) from BSPM. We found that the feature corresponding to CLs of 220 – 230 ms had the highest hazard ratio on



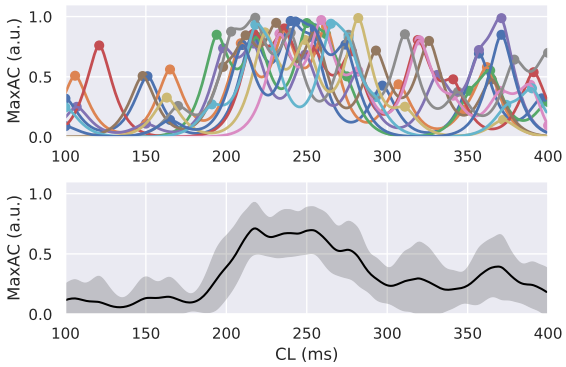
(a) Patient 1 ($1 \leq l \leq 5$)



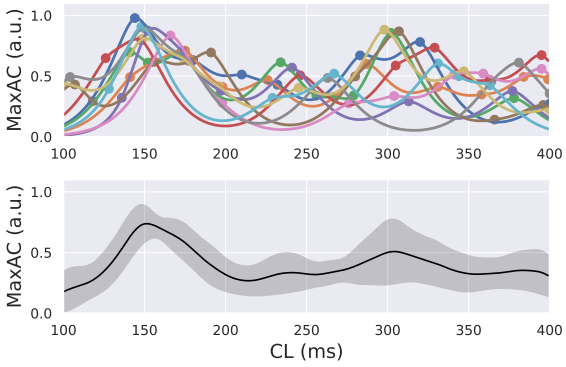
(b) Patient 2 ($1 \leq l \leq 5$)



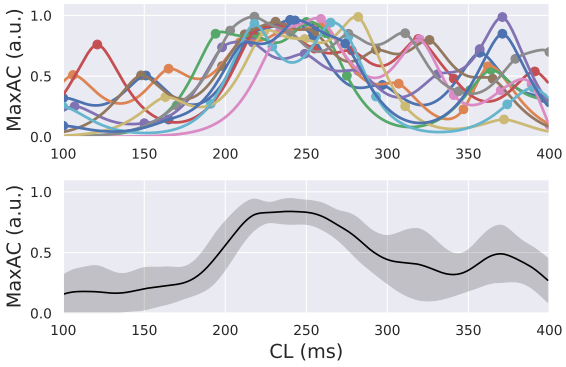
(c) Patient 1 ($1 \leq l \leq 10$)



(d) Patient 2 ($1 \leq l \leq 10$)

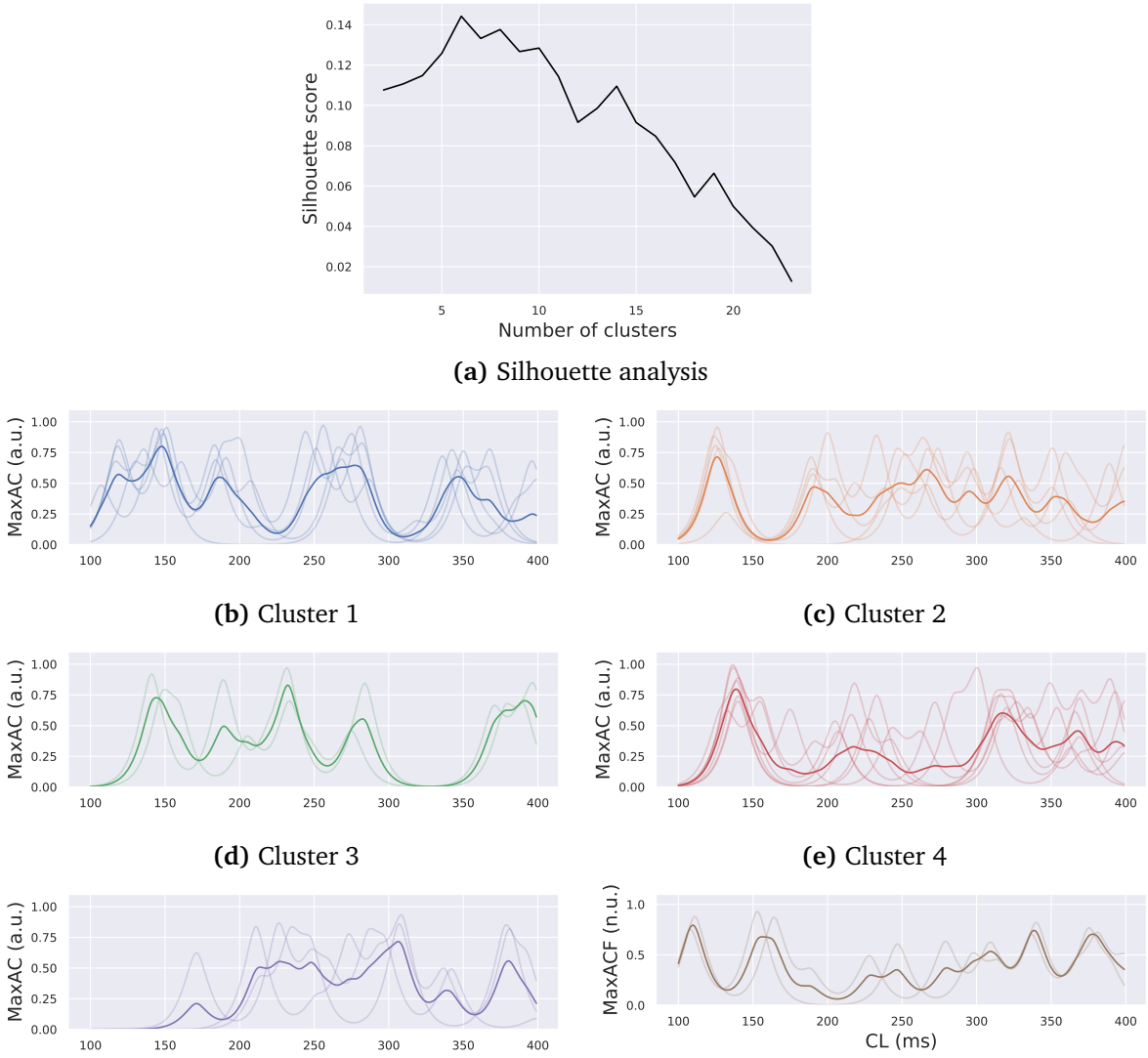


(e) Patient 1 ($1 \leq l \leq 20$)



(f) Patient 2 ($1 \leq l \leq 20$)

Fig. 5.11. Examples of non-invasive APSS from BSPMs of Patient 1 (a,c,e) with a peak-shape pattern, and Patient 2 (b,d,f) with a platform-shape pattern, estimated from BSPM with GPR (with $1 \leq l \leq 10$). In each panel, the top plot shows the non-invasive APSS curves derived from an f-wave of the patient, and the bottom plot shows the mean (black plot) and standard deviation (gray shade) of the APSS curves at the top plot. a.u.: arbitrary unit.



(a) Silhouette analysis

(b) Cluster 1

(c) Cluster 2

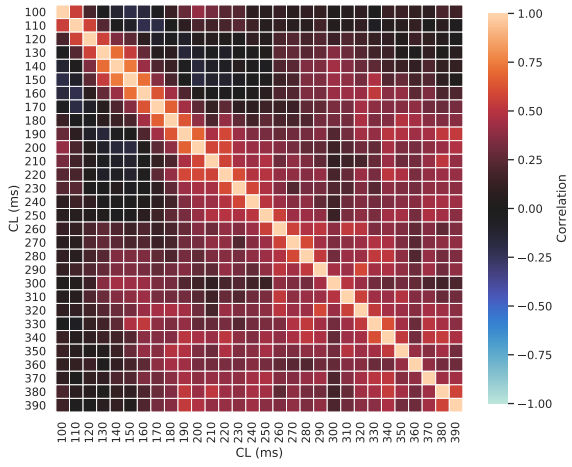
(d) Cluster 3

(e) Cluster 4

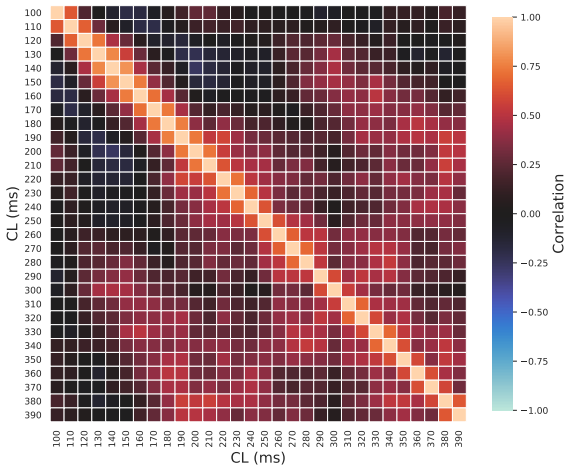
(f) Cluster 5

(g) Cluster 6

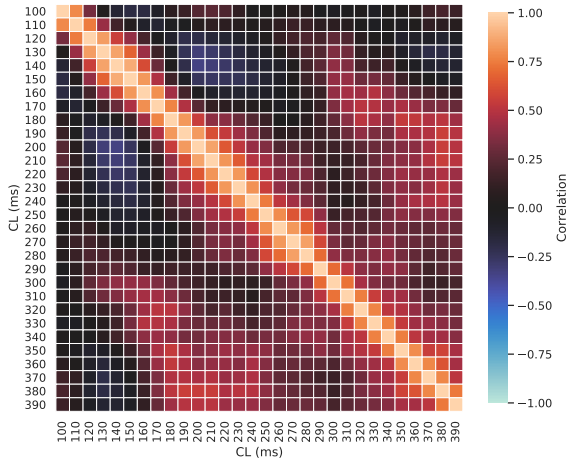
Fig. 5.12. The non-invasive APSS curves of a patient, estimated from BSPM with GPR (with $1 \leq l \leq 10$). (a) The silhouette score on different number of clusters, which picked 6 clusters for the K-means clustering method. (b-g) The non-invasive APSS curves of each cluster, where the thick curve in each plot shows the mean shape of the curves. a.u.: arbitrary unit.



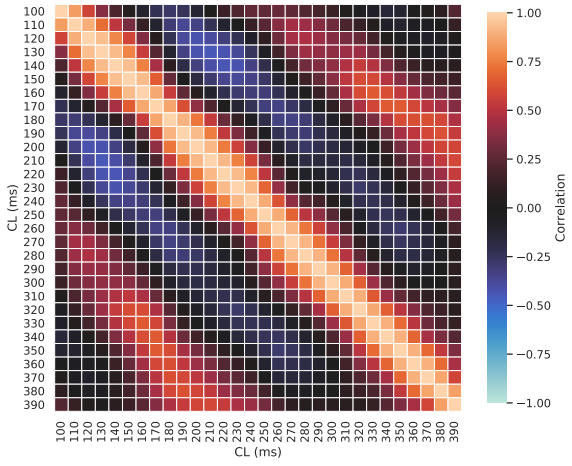
(a) Non-invasive APSS ($1 \leq l \leq 5$)



(b) Non-invasive APSS ($1 \leq l \leq 10$)



(c) Non-invasive APSS ($1 \leq l \leq 20$)



(d) MaxAC without SO-BSS

Fig. 5.13. The correlation between features on the persistent AF dataset, for (a-c) non-invasive APSS estimated from BSPM with SO-BSS and GPR, and (d) MaxAC directly on BSPM without SO-BSS.

all training datasets of 10-fold cross-validation splits. The evaluation on the final assembled test set also showed that the high-value group was significantly better than the low-value group with a p-value of 0.006 in the AF-free rate under a logrank test [214], as shown in Fig. 5.15. Examples of patient non-invasive APSS estimated with SO-BSS and GPR ($1 \leq l \leq 10$) from BSPM, of high and low MaxAC values at CLs of 220 – 230 ms, are shown in Fig. 5.16, with different morphologies of APSS curves.

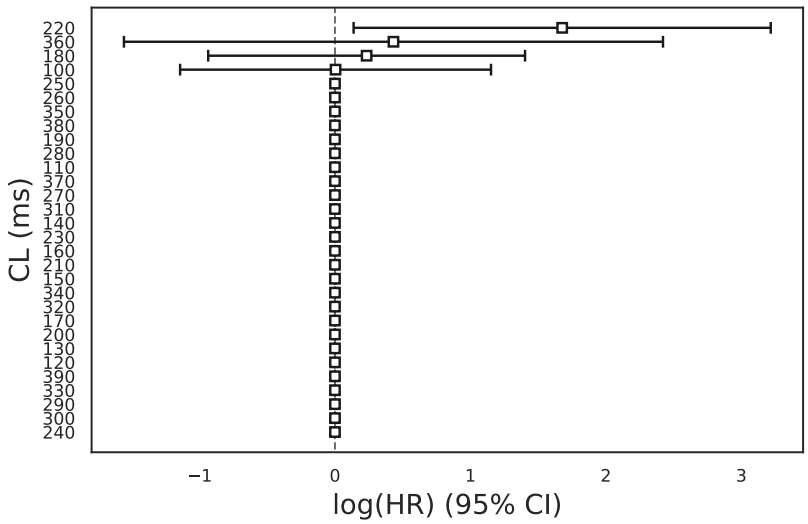
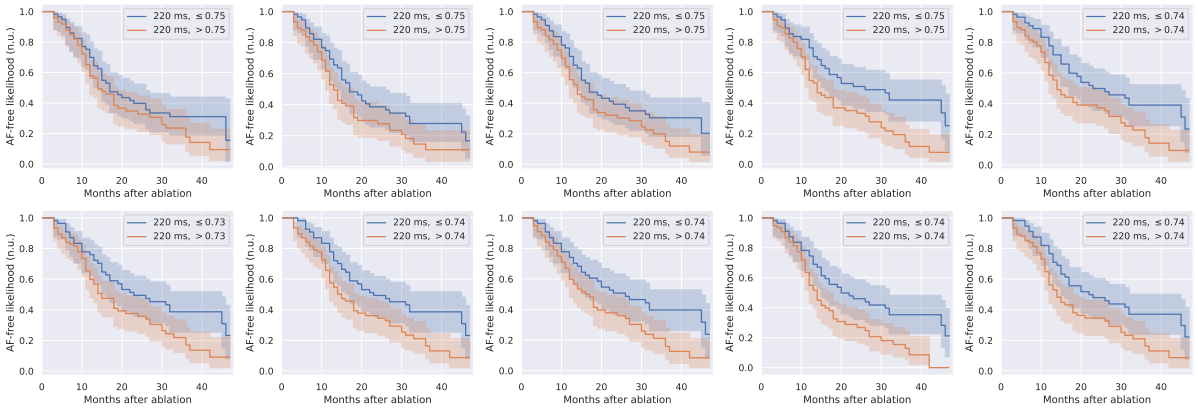


Fig. 5.14. Log hazard ratios with 95% confidence interval (x-axis) of all features derived from the non-invasive APSS from BSPM (estimated with $1 \leq l \leq 10$), estimated by Cox proportional hazard model with L1 regulation. A hazard ratio >1 means a higher value of the factor results in a decrease in the AF-free likelihood (survival rate), compared to baseline AF-free likelihood. The feature "220 ms" representing the patient MaxAC content on CLs of 220–230ms has a hazard ratio > 1 with 95% confidence.

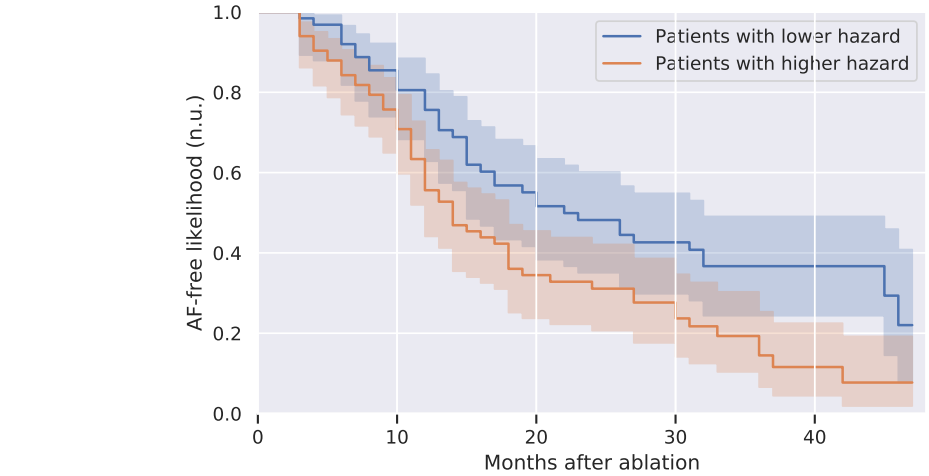
We also tested this on the non-invasive APSS calculated from BSPM and ECG, without and with SO-BSS, and show their p-values in Table 5.1. Except for non-invasive APSS calculated from ECG without SO-BSS, non-invasive APSSs were able to predict two patient groups with different outcomes. The smaller bound is associated with a better prediction performance, with smaller p-values in the logrank test for $1 \leq l \leq 5$ and $1 \leq l \leq 10$. Features of 370 – 400 ms and 210 – 240 ms produced high hazard ratios.

5.4 Discussion

In this study, we presented non-invasive APSS from either BSPM or ECG, a novel representation of patient AF condition based on the atrial periodic sources over a wide range of CLs. Although the APSS is similar to spectral analysis, it has a finer time resolution, the same with the sampling frequency, much higher than the FFT analysis on



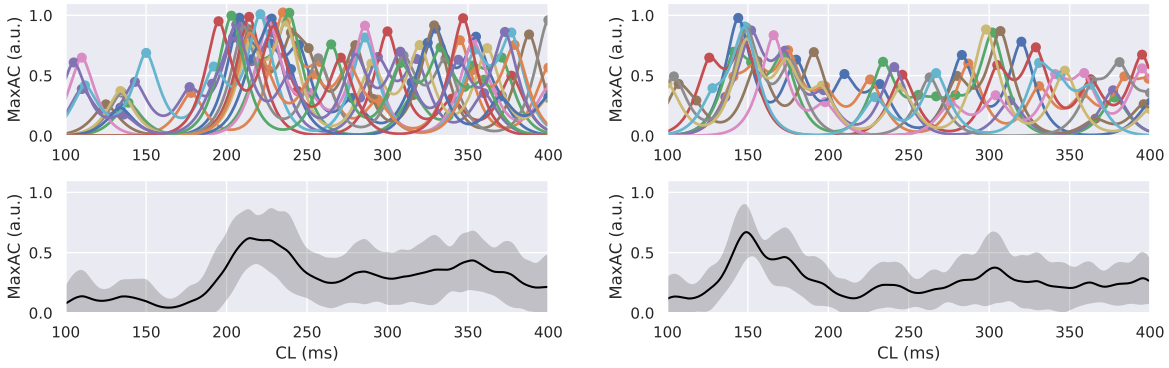
(a)



Patients with higher hazard					
At risk	83	57	21	12	3
Censored	1	3	13	16	20
Events	0	24	50	56	61
Patients with lower hazard					
At risk	63	49	30	23	8
Censored	0	2	4	6	18
Events	0	12	29	34	37

(b)

Fig. 5.15. 10-fold cross-validation Kaplan-Meier plots for feature selection of APSS estimated from BSPM with SO-BSS and GPR ($1 \leq l \leq 10$), with a Cox's proportional hazard LASSO model, on the persistent AF datasets. (a) Train set in each of the 10 folds. (b) Evaluation on the final assembled test set, with p = 0.006 in the logrank test. In each fold, a single feature with the highest hazard ratio and the decision boundary is selected from the training fold. In the legend of (b), at a particular time shown vertically above, "At risk" and "events" show the count of patients without and with AF recurrence up to that time, respectively, and "censored" shows the number of patients who did not have AF recurrence up to that time, but no event was recorded afterwards.



(a) **(b)**

Fig. 5.16. Example patterns of non-invasive APSS estimated from BSPM with SO-BSS and GPR ($1 \leq l \leq 10$), with (a) high and (b) low MaxACs at CLs of 220 – 230 ms. The legend is the same as Fig. 5.11. Patient of (a) experienced AF recurrence at 9 months after ablation, and patient of (b) experienced no recurrence up to 37 months post-ablation. a.u.: arbitrary unit.

Table 5.1. P-values for logrank tests with the null hypothesis that the two predicted patient groups have the same survival likelihood, using features from non-invasive APSS extracted from different sources (BSPM and ECG), and by different methods (estimated with SO-BSS and GPR or computed directly). For the entries with [1], features of 370–380 ms, 380–390 ms or 390–400 ms were found to be the highest hazard; other entries identified features of 210–220 ms, 220–230 ms or 230–240 ms with the highest hazard ratio. The bold font with two stars highlights entries with p-values < 0.01, and with one star show entries with p-values < 0.05. l is the time-scale of the Matérn kernel in the GPR.

	Estimated with SO-BSS and GPR				Computed directly
	$1 \leq l \leq 5$	$1 \leq l \leq 10$	$1 \leq l \leq 20$	$1 \leq l \leq 30$	
From BSPM	**0.004****	**0.006****	0.07	**0.04***	**0.01***[1]
From ECG	**0.003****[1]	**0.03***[1]	0.08[1]	0.15[1]	0.6[1]

short signals. As a multi-lead analysis, our prediction of dominant CL is also less likely to result in an underestimation of the DF presented in FFT analysis [136]. They reveal patient-specific atrial periodic wave patterns, and were able to predict patient groups with an improved long-term outcome. The APSS can also be derived from the intracardiac recordings. This provides a way to validate our non-invasive APSS, and also shows its usefulness for a variety of applications, ranging from intra-operative to home use.

5.4.1 Technical considerations of non-invasive APSS

We showed the importance of SO-BSS in estimating APSS non-invasively. Usually, frequency-domain analysis such as the FFT is used for extraction of periodic components. However, the length of the signal should be short enough for analyzing AF from a f-wave of a single beat, rendering a low time precision of the periodic components by the FFT (Fig. 5.6(b)). Using MaxAC as an empirical measurement of the periodic content is another choice, but directly applying SO-BSS to BSPM loses some periodic content, as presented in Fig. 5.6(a) and Fig. 5.7(d-e). For estimating the APSS, SO-BSS is especially helpful for ECG analysis, as ECG has fewer leads, which was likely the reason that using ECG only without SO-BSS failed to predict patient groups with distinctive outcomes in Table 5.1.

The non-invasive APSS is based on the CL-MaxAC pairs and modeled by an additional Bayesian model GPR. We presented the technical consideration for choosing the hyperparameter l for the Matérn covariance function of GPR, and found that $1 \leq l \leq 10$ achieved a good balance of revealing patient-specific APSS patterns (Fig. 5.9), and was useful for the survival prediction for both BSPM and ECG (Table 5.1). The patient-specific pattern(s) between non-invasive APSS from different episodes of persistent AF patients showed evidence that common patient-specific patterns of APSS can be extracted from AF signals of high complexity for various patients, and also validated our hypothesis that the periodic components can represent the AF condition of a patient.

Our non-invasive APSS approximated the true APSS with a small MAE and a high coverage, although some error was still unavoidable due to the intrinsic volume-conductor nature of far-field signals, reducing the amplitude of the MaxAC. Besides, we used a computationally tractable way to calculate the samples drawn from the true APSS, by extracting the mean shapes of similar transmembrane voltages with unsupervised learning, which could result in an underestimation of the MaxAC of true APSS. Validation to compare between the intracardiac electrograms and torso potentials is needed. The CL estimated by our method also did not have a systematic bias (under- or over-estimation), as shown in Fig. 5.10, but further validation on complex AF episodes, where the atrial activities are less organized, is needed.

Although one may consider using dimension reduction techniques such as Principal Component Analysis [16], to reduce the number of features with a strong correlation, we deliberately kept the simplest classification model and a uni-variate feature for predicting patient groups with different outcomes, for the clarity of its physiological meaning. We also chose to focus on markers extracted solely from body surface potential signals, to explore the potential predictive values of such signals, but combining other bio-markers of the patients would likely improve the survival prediction.

5.4.2 Mechanistic insights from survival analysis with non-invasive APSS

Using the Cox proportional hazard model with L1 regularization, MaxAC on the CL intervals of 210 – 230 ms and 360 – 400 ms were identified as two simple uni-variate predictive metrics on our cohort of persistent AF patients. According to the correlation plot in Fig. 5.13(c), we relaxed these CL ranges to their adjacent CL that are moderately correlated (with a Pearson coefficient around 0.5), which gave the two frequency bands of 4 – 5 Hz (200 – 250 ms) and 2.5 – 2.9 Hz (350 – 400 ms) that indicated high hazards for post-ablation AF recurrence in our datasets.

In particular, periodic sources of 4 – 5 Hz and 2.5 – 2.9 Hz are likely to be slow reentries with large conduction blocks, or macro-reentries (Fig. 6.1), as the frequency of rotors was typically ≥ 5 Hz in human AF [12]. There are several possible explanations. Reentries with anatomical conduction blocks or macro-reentries, are less straightforward to map by a catheter, and therefore, could be missed by traditional strategies such as PV isolation and rotor ablation. As a long AF CL of 190 ms already indicates a higher degree of organization in the atria [114], periodic sources with CL above 200 ms may be mistaken as satisfactory immediate outcomes for termination. Furthermore, as the frequency and the amplitude of the highest DFs are usually adopted for analyzing AF, coexisting periodic sources which are not the highest DF could be missed. All these considerations are conjectures from literature, and so require verification using intra-cardiac signals of AF patients in prospective studies.

5.4.3 Limitation

We constructed APSS using isotropic and stationary kernels. The isotropic kernel assumes that the diminishing speeds of the MaxAC by increasing and decreasing CL are the same, whereas the stationary kernel assumes that the diminishing speeds are the same for all CLs. This is likely an over-simplification, and a fine-tuning of the kernel functions should be considered with validation using experimental or patient data.

Chapter 6

Non-invasive detection of active and passive centrifugal activation with deep learning

6.1 Introduction

Centrifugal activation (CAT), which refers to a focal activation starting at a central site and propagating radially to the surrounding tissue, is an electrophysiologically important wave propagation pattern. It presents at the activation of pacemaker cells, the initiation of ectopic beats, and when electrical propagation travels through inter-atrial connections as passive CATs, including the interatrial septa (fossa ovalis and Bachmann's bundle) and the coronary sinus ostium, all confirmed by high-resolution mapping on 35 AF patients in a recent multi-center study led by the the Haut-Lévèque Cardiology Hospital of Bordeaux [215]. They concluded that the differentiation between active and passive CATs helps to identify the true arrhythmic mechanism[215].

We further hypothesized that a passive CAT, when regularly presenting on a single atrial chamber, is a consequence of stationary driver(s) on the opposite atrial chamber, as the whole atria is entrained by the dominant rhythm of the driver. Non-invasive mapping of spatially stationary drivers can help narrow down the mapping scope for invasive mapping.

Furthermore, the occurrence of passive CATs on two atrial chambers alternatively and regularly may indicate an inter-atrial macro-reentrant circuit (Fig. 6.1), as they are uni-directional and regular. Non-invasive identification of inter-atrial macro-reentrant circuits further suggests ablation sites, such as coronary sinus ostium, which was found to be an important isthmus for an inter-atrial macro-reentrant circuit [216]. An in-silico study [165] also suggested the ablation of inter-atrial connections could be useful

for some persistent AF patients, depending on the geometry of the AF. However, it is also particularly challenging to detect the macro-reentry accurately without entrainment mapping used by the clinicians, with only about 50% accuracy using a directed graph approach from invasive high-density activation mapping [217].

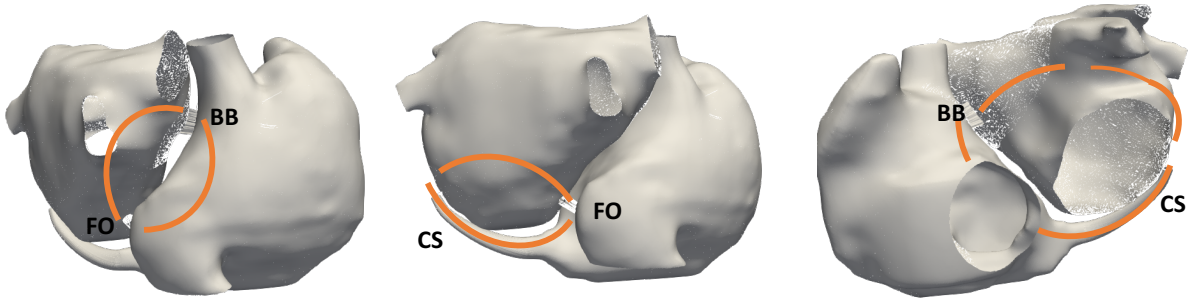


Fig. 6.1. Illustration of inter-atrial macro-reentrant circuits via the inter-atrial connections of BB, CS and FO, with the macro-reentrant circuit paths shown in orange. The passive CAT can occur at the junctions. BB: Bachmann's bundle. CS: coronary sinus. FO: fossa ovalis. Note that this does not correspond to the travelling direction of the activation wavefronts.

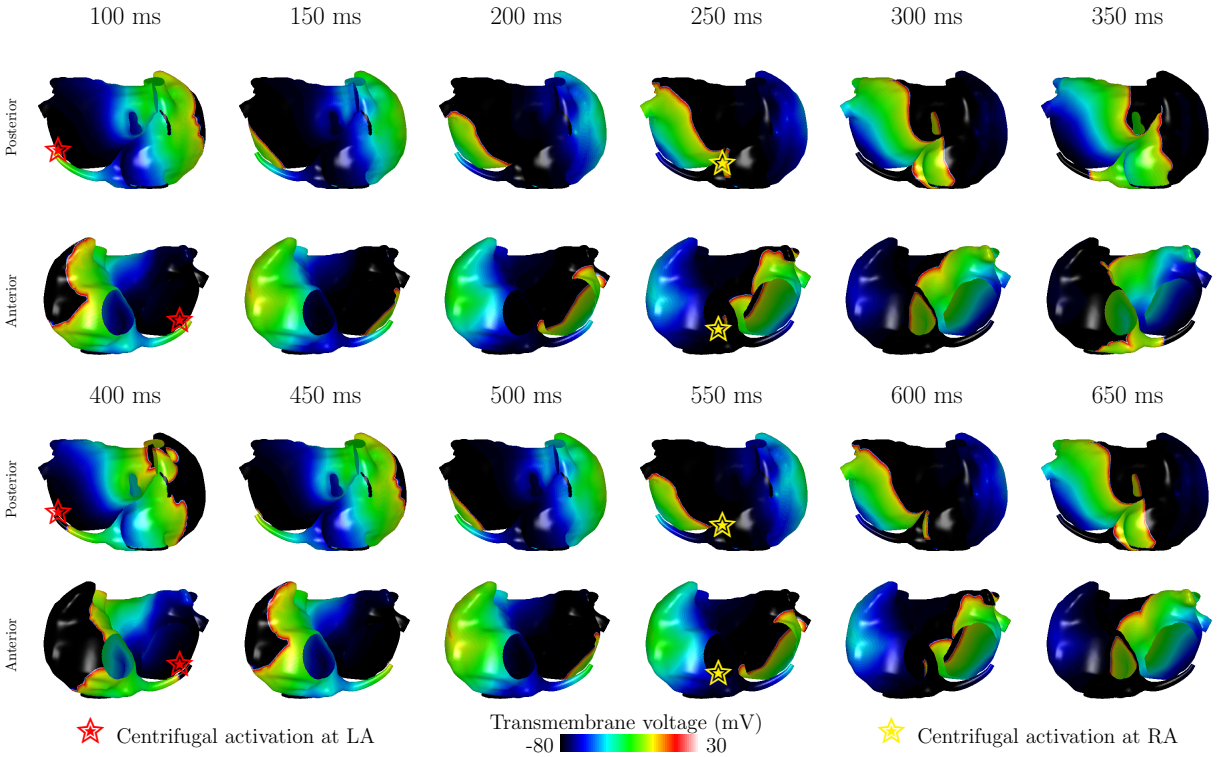


Fig. 6.2. An example of an intra-cardiac macro-reentry circulating via the coronary sinus connection with the LA, and the fossa ovalis with the RA, with a CL of 370 ms.

ECGi has sometimes been used to identify macro-reentries involving both atria. Cuculich et al. [40] identified a slow bi-atrial macro-reentry with a rotor core from a reconstructed phase map by ECGi, on a low-resolution bi-atrial mesh without any inter-atrial connection structure. This may miss intra-atrial macro-reentries without a rotor

core. Such case includes an inter-atrial macro-reentry circulating between the coronary sinus and the fossa ovalis as in Fig. 6.2, and an inter-atrial macro-reentry circulating between the Bachmann's bundle and the fossa ovalis septa as in Fig. 6.3, which has a shorter macro-reentrant pathway, and thereby, a comparatively faster speed than going through coronary sinus. Detection of macro-reentry via spectral analysis is also not feasible. As mentioned in Chapter 4, using SO-BSS with a random forest classifier, macro-reentries could also be detected as FS falsely, as the extracted periodic sources of macro-reentries are also highly periodic and synchronized.

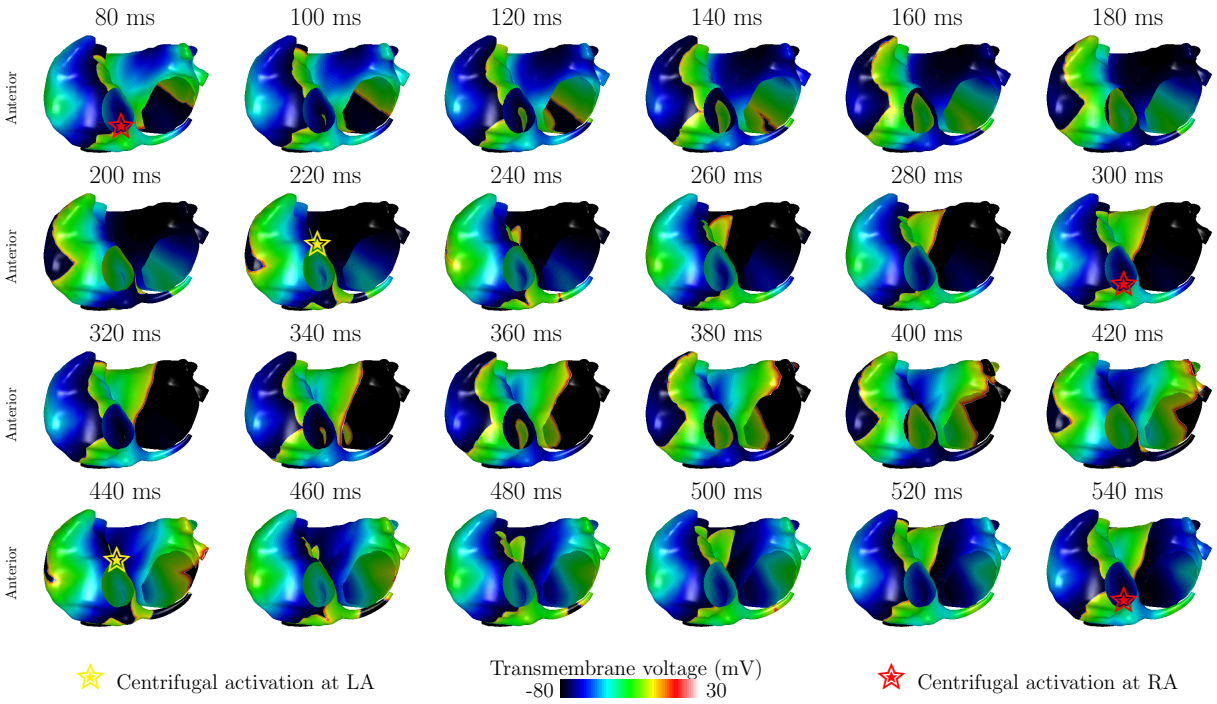


Fig. 6.3. An example of an inter-cardiac macro-reentry circulating via Bachmann's bundle on LA and fossa ovalis on RA with a cycle length of 223 ms.

Supervised machine learning can learn the correspondence between input data and prediction target via training examples. For some inputs, especially high dimensional, an additional step of feature extraction is required. However, it can be more effective to use deep learning to recognize the morphology of electrophysiological signals rather than using hand-crafted features [218]. The non-invasive extraction of timing and location relies on exploiting the spatial-temporal structure of BSPM, which can be viewed as a time series of 252-lead potentials, similar to a video consisting of a series of images, where the locations of the image pixels are irregular. We first considered representing BSPM as a convolutional recurrent neural network (CRNN) architecture, which applies convolutional filters such as convolution, pooling, and batch normalization to encode each frame in a flattened vector, and learns the temporal relationship of this encoded representation through a recurrent neural network (RNN). In particular, the weights in the cells of RNN were shared, so that common features could be learned from different parts of a f-wave, making the RNN suitable to encode the cyclic structures in the repet-

itive focal firings and reentrant waves during AF. CRNN has been extensively applied to encode spatial-temporal data such as videos [219]. In cardiology, it has also been adopted successfully for end-to-end event detection in medical imaging, such as end-diastole detection from ultrasound [220], and the diagnosis of cardiac abnormalities from ECG, such as [221].

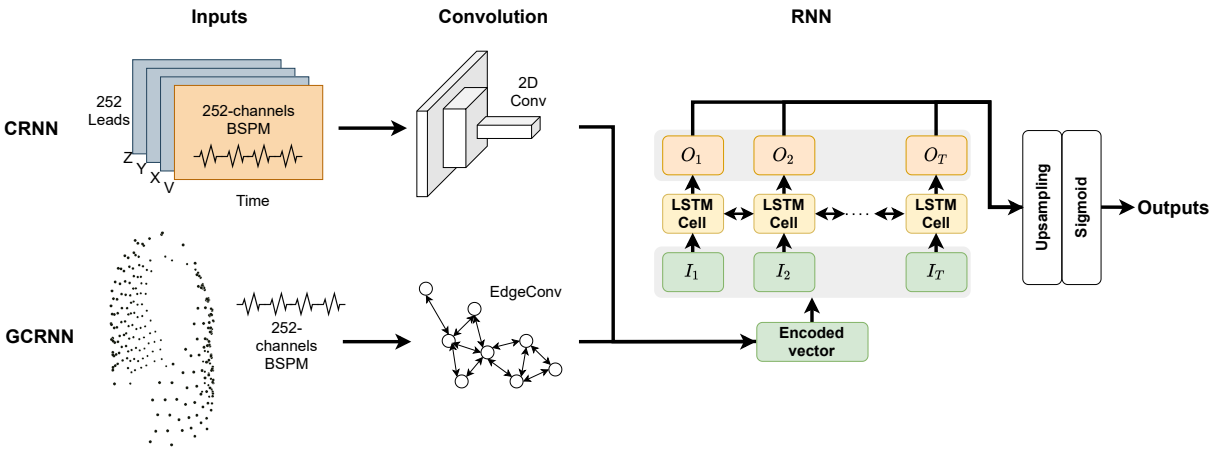
However, the irregular arrangement of the BSPM electrodes poses difficulty in applying the traditional convolutional operations, which are used on regular pixels or grids. To solve this, we propose that using a graph to represent BSPM leads, and a generalized convolution operation for non-euclidean graphical neural network (GNN) models may be more appropriate, as the BSPM leads can be represented as a graph of nodes where the neighbouring nodes share similar information from the same cardiac sources. We proposed a Graph Convolutional Recurrent Neural Network (GCRNN), combining an RNN with a GNN, where the GNN replaces the 2D convolution component in the CRNN with an edge convolution operator (EdgeConv) [99] to perform convolution operations from the nearest neighbours of each electrode.

The goal of this study was to develop a suitable deep learning model for non-invasive detection of 1) active CATs, 2) the origins of FS drivers, 3) passive CATs, and 4) the atrial chamber which the passive CATs land on, from BSPM recordings. We attempted to use a GCRNN model, combining a GNN and a RNN. Its performance was compared to a baseline CRNN, consisting of a standard CNN and a RNN.

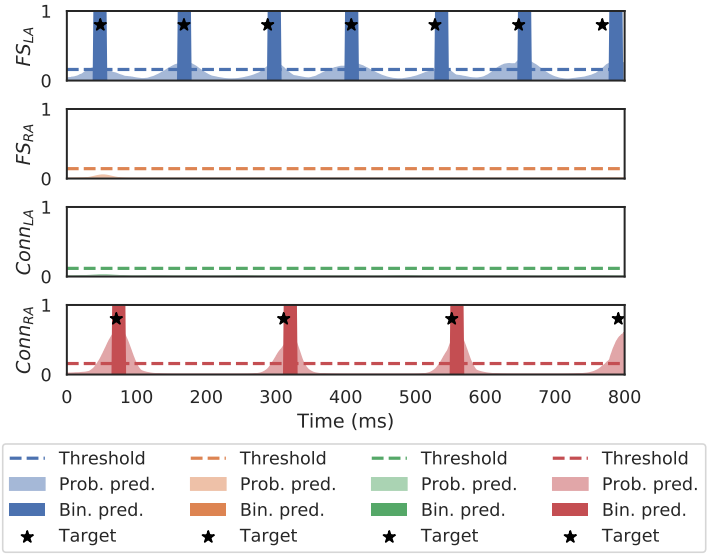
6.2 Methods

6.2.1 Graph Convolutional Recurrent Neural Network

The architecture of the GCRNN and an example output of our method are shown in Fig.6.4, on a 252-channel T-length BSPM signal. On the left, there are two alternative convolution layers to transform input BSPM data into a condensed spatiotemporal tensor of size $252 \times T_c$, where T_c is the time length of the compact temporal vector with a reduced time length and $T_c < T$. This constitutes two alternative approaches, a traditional 2D convolution (2D Conv) layer, and an Edge Convolutional layer (EdgeConv), corresponding to the CRNN and GCRNN, respectively. In the middle, an RNN learns the temporal dependency of the condensed representation from a series of the compact BSPM embedding. The output is a function over time which indicates the probability of CAT as classified by the site (FS_{LA}, FS_{LA}, $Conn_{LA}$, and $Conn_{RA}$). With a cut-off threshold obtained by cross-validation, the probabilistic output function can be transformed into an output function with only binary values.



(a) Architectures of CRNN and GCRNN



(b) Probabilistic and binary outputs for four classes, as well as network targets.

Fig. 6.4. (a) Architectures of CRNN and GCRNN, which output (b) the time-dependent probabilistic functions of $CAT(t)$ for FS_{LA}, FS_{RA}, $Conn_{LA}$, and $Conn_{RA}$. The description of the architecture is presented in the main text. The example output is from an FS on the LA of 120 ms. The probabilistic prediction (prob. pred.) is represented as the half-transparent shaded area. The cut-off threshold is shown by the dash lines. The binary prediction (bin. pred.) is represented as the vertical bars with zero opacity. The network prediction targets are marked by black stars. In (a), for the input of the CRNN, a 252-channel T-length BSPM signal was represented as $4\times252\times T$ tensor of value (V), and three-dimensional coordinates of the vest leads (X, Y, and Z), whereas for the input of a GCRNN, the 252-channel BSPM was represented as a network of 252 nodes, with signal of the electrode being the feature of that node.

GNN as a spatiotemporal encoder for BSPM We represented the 252 lead positions as network nodes, and their signals of length T as features of the node. The EdgeConv operator updates the features of each network node by learning another feature, to

compact the T-length features into a feature of length $T_c = 100$, by passing update messages through neighbour nodes. The compacted signals of each node make the training of the RNN easier. Specifically, the EdgeConv operator sums up the information propagated from the neighbouring nodes, where the information is learned by a simple multi-layer perceptron (MLP) network. Formally, the operation of EdgeConv is defined by the following equation:

$$\mathbf{x}_i = \sum_{j \in \mathcal{N}(i)} MLP(\mathbf{x}_i || \mathbf{x}_j - \mathbf{x}_i)$$

where the neighbours $\mathcal{N}(i)$ are defined by the $N_{neighbour}$ nearest neighbours of each BSPM node given the three-dimensional locations of the leads in the Euclidean space. We chose $N_{neighbour} = 6$, and used $||$ to denote the concatenation of two tensors by expanding the number of features, as the BSPM leads were designed as clusters of six nodes surrounding one node (Fig. 2.6). The MLP network consists of a fully connected layer that maps the concatenation of the vector and the difference with its neighbouring node ($\mathbf{x}_i || \mathbf{x}_j - \mathbf{x}_i$), to a T_c-length tensor, where T_c is the length of the signal. Afterwards, a ReLU layer maps this T_c-length tensor with 252 input features to another tensor of the same shape using a fully connected layer.

RNN as a temporal encoder An RNN with was then used to extract the temporal structure from the compacted representation learned from the GNN or CNN. An RNN learns the common patterns of adjacent time steps from the entire time series. It consists of a chain of cells, and the representation of each time step of a time series was taken as the input of each cell. A learned representation coming out from each cell is fed back to the cell recurrently, with additional input from a $(t+1)$-th input. We considered an LSTM cell [92], which can learn a temporal encoding over a long sequence through gating functions on variables in the cell. To enable the learning of the pattern in both the forward and backward directions, we used a bidirectional LSTM. This means that the information can flow in both directions of the times series. The output of each cell is then an embedding of the input time series, the length of which is the same as the input time series. Readers interested in RNN and LSTM are referred to Chapter 2 for their formulae. A LSTM cell had an input size of 252.

Upsampling layer The output of the RNN is then transformed by a dynamic upsampling layer to the original time length. The upsampling operation is a transpose of the convolution operation, which upsamples each entry of a vector to its multiple. The parameters of this layer are dynamically updated during the training phase of the network [222]. Similar to the weight sharing of a convolutional layer, the weights of the upsampling layer are also shared across different patches of the whole time series.

Output layer The network then goes through a sigmoid function to map the output between 0 and 1. The final output vector is a long probability vector of the same number of classes as the time length, 800 classes in our case, similar to the network output in our previous work [218] annotating the repolarization time on filter-distorted unipolar electrograms. This output vector then goes through a moving average filter of size 20 ms, to smooth out any spurious spikes.

Prediction target The prediction target is a vector of the same length as the time series, where each entry of the vector corresponds to the likelihood of the onset of the CAT of a time instance. To detect the onset of the CAT, we marked entries of the prediction output vector corresponding to the first 20 ms since the onset of the CAT as 1, and marked the other entries as 0.

Neural network training

Similar to Chapter 4, we used a leave-one-patient-out cross-validation method to perform the validation and evaluation process. We arranged the five patient simulations by taking three patients for training, one patient for the validation set, and the trained model was evaluated on the remaining holdout patient. We varied the holdout set to be all five different patients.

To augment the training dataset for building more robust networks, we started with an 800 ms segment and took five subsets of 750 ms with 0, 10, ... 50 ms as the starting point during training. During the evaluation, these six subsets were aggregated by averaging to form one 800-ms segment. To train the model, an Adam optimizer [223] was used with a learning rate of 0.0001. A binary cross entropy loss was used to evaluate the probabilistic outputs of both training and validation sets. The total binary cross entropy loss was defined as the average of sample binary cross entropy loss. The sample binary cross entropy loss l of the i-th sample for the c-th class, $p_{i,c} \in [0, 1]$, evaluated on its binary ground truth, $y_{i,c} \in \{0, 1\}$, was defined as

$$l = -\left[y_{i,c} \cdot \log p_{i,c} + (1 - y_{i,c}) \cdot \log(1 - p_{i,c})\right] \tag{6.1}$$

The network was updated by going through the whole training dataset using mini-batch gradient descent with a batch size of 32. Each pass going through the whole training dataset was marked as an epoch, where the binary cross-entropy loss of the validation was computed at each epoch. The training stopped after seeing the validation loss continuously increasing for 20 epochs. The network at the "best" epoch, i.e. when the validation loss was smallest, was adopted as the final prediction model.

The validation set was used to select the best epoch and a cut-off threshold to

obtain a binary output from the probabilistic output from the trained model. The cut-off threshold of a validation set was obtained from the threshold resulting in the highest Area under Receiver-Operator curve (AUROC) of the probabilistic outputs on predicting the presence of a CAT over the validation set. The final prediction for CAT function used the maximum of predicted CAT function of each time step predicted over all the train-validation splits, and the final cut-off threshold was the largest cut-off threshold across the validation sets.

To further minimize the impact of spurious outputs on the extraction of the binary outputs, we extracted peaks of the CAT functions exceeding the cut-off threshold, while ensuring the distance between two detected peaks to be $>$ 100 ms by preserving only the highest peaks. Intervals starting from 10 ms before until 10 ms after the time of CAT peak were marked as the timing of CAT onset, where the length of 20 ms was set to be the same as the prediction target.

The network was trained using PyTorch version 1.4, torch-geometric version 1.6.3, CUDA version 10.2 on a Quadro RTX 8000 Graphic Processing Unit. It was faster to train the GCRNN than the CRNN, as it took 20 – 30 minutes to train 27 – 41 epochs for the GCRNN, while it took 40 – 60 minutes on 23 – 25 epochs for the CRNN, when training the networks on a three-patient training dataset and a one-patient validation dataset (Patient 5).

6.2.2 Preparation of prediction inputs and targets

This project used the synthetic dataset built in Chapter 3. In summary, the dataset contains 1772 (driver-type FS) + 328 (initiator-type FS) + 328 (FS-induced AF) + 549 (reentrant AF) = 2977 f-wave segments.

Computation of CAT

On the atria, we computed the timing of the active CAT as the origin of the FS, denoted as FS_{LA} or FS_{RA}, as well as the passive CATs, which arrive at the opposite atrial chamber through inter-atrial connections, denoted as $Conn_{LA}$ or $Conn_{RA}$. The local activation time of the surrounding tissue was calculated as the duration between the activation in the node since the first activation at the center of CAT. We considered local activation time only within 10 – 200 ms from the activation of the candidate focal site. To avoid inaccuracy, we only labeled and predicted CATs up to $1000 - 200 = 800$ ms in our signals. We then labeled the presence of CAT using the criterion that the distances from the centroid were linearly correlated with their activation times since the activation of the centroid, measured by a correlation coefficient of > 0.8 by a Pearson correlation test. The process is illustrated in the Fig. 6.5.

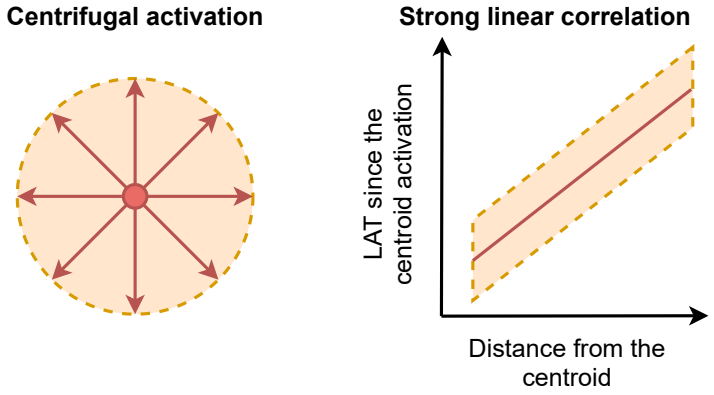


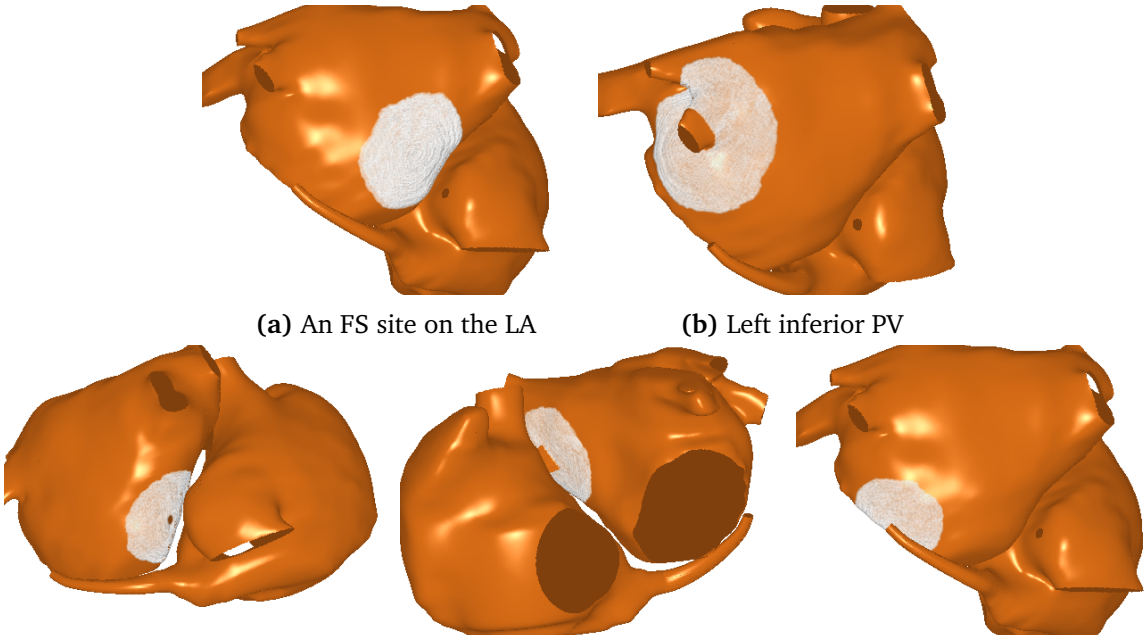
Fig. 6.5. Detection method for the CAT. Local Activation Time: Local activation time. The strong linear correlation is tested by the Pearson correlation test with a correlation coefficient > 0.8.

From our experience, a CAT is only detectable from the BSPM when the radius of CAT is large enough, although this boundary was unclear to us. For example, for those fast FS which induce a 2:1 activation to the surrounding tissue, only a half of the CATs were detectable. Instead of labelling the presence of focal sources by prescribed labels, we only included those CATs propagating through a detecting radius, without blocks, as the detection targets. The annotation of CAT considers those radial activations emanating for at least 50, 100, and 150 triangular edges from the centrifugal centers, which correspond to approximately 2 cm, 4 cm and 6 cm away from the activation centers. These radii were loosely selected, and could be optimized in future projects. To demonstrate the radial activations emanating for at least 50 triangular edges from the centrifugal centers, we showed the area with ≤ 50 triangular edges from the centroid in Fig. 6.6. Compared to using the Euclidean distance in the three-dimensional space to find such regions, this method of extraction by the number of connected edges did not include sites on the opposite chamber (Fig. 6.6 (c, d)) or on the coronary sinus (Fig. 6.6 (e)).

AF mechanism prediction with CAT

The CAT occurrences of FS_{LA}, FS_{RA}, $Conn_{LA}$ and $Conn_{RA}$ over 800 ms were used to infer the AF mechanisms. We detected FS on LA and on RA, which were represented by ≥ 3 occurrence of FS_{LA} for LA FS, and ≥ 2 occurrence of FS_{RA} for RA FS.

From $Conn_{LA}$ and $Conn_{RA}$, the presence of inter-atrial macro-reentrant circuit can be extracted by if there are repetitive CATs on both $Conn_{LA}$ and $Conn_{RA}$. Here, an alternating occurrence of the $Conn_{LA}$ and $Conn_{RA}$ would be more accurate, as activity follows a biatrial conduction pathway via two inter-atrial connections. However, this requires high accuracy in estimating the timings of $Conn_{LA}$ and $Conn_{RA}$, and therefore, we relaxed the criterion of alternation, but only counted the occurrences of $Conn_{LA}$ and



(a) An FS site on the LA **(b)** Left inferior PV

(c) Fossa ovalis on the LA **(d)** Bachmann's bundle on the LA **(e)** Coronary sinus on the LA

Fig. 6.6. Illustration of the minimal area (in white) for labelling CAT targets, with $\leq$ 50 triangular edges from (a) an LA site at ($\alpha_{LA} = 0.2, \alpha_{RA} = 0.2$) (b) left inferior PV, and (c-e) inter-atrial connections on the LA.

$Conn_{RA}$ to be ≥ 2 times each as the criterion of inter-atrial macro-reentries.

If there was no macro-reentrant circuit, the entries with ≥ 3 repetitive $Conn_{RA}$ and ≤ 1 $Conn_{LA}$ were tagged as "driver on LA". These drivers emitted repetitive atrial propagation in one direction from the LA to the RA, causing a repetitive activation at the inter-atrial connection with the RA, and hence, $Conn_{RA}$ was repetitively activated. Meanwhile, wave breaks at RA caused only occasional activation of $Conn_{LA}$. Likewise, we tagged entries with ≥ 2 repetitive $Conn_{LA}$ and ≤ 1 $Conn_{RA}$ as "driver on RA", which promotes repetitive CAT arriving at the other atrial chamber through inter-atrial connections. The required number for $Conn_{LA}$ decreased for the reason that RA conduction is slower [34].

6.3 Results

We first discussed the validation of using CAT as prediction targets to infer AF mechanisms. Afterwards, we evaluated our network performance by the CAT timing as well as the AF mechanism, comparing with a vanilla CRNN, the architectural details of which can be found in Appendix B.

We evaluated the prediction of AF mechanisms in the form of a multi-class matrix A_{ij}, where the subscripts ij shows the number of a targeted class i (in a row) was

predicted as class j (in a column). When the matrix is normalized by the number of targets in each class, the entries on the diagonal show the recall scores of the true classes, and when normalized by the number of predictions in each class, the diagonal shows the precision scores of the true classes.

6.3.1 Validity of using CAT as a prediction target to infer AF mechanisms

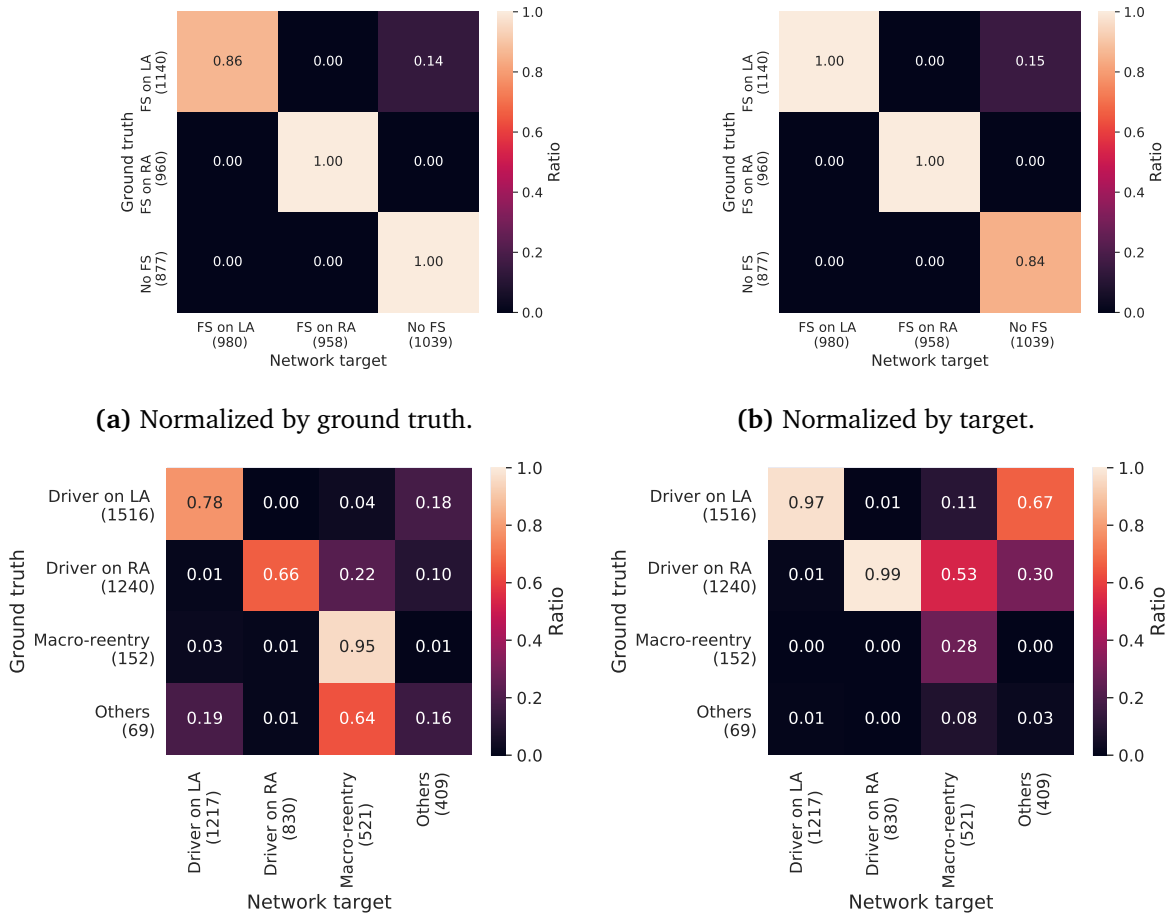
We assessed the suitability of using CAT as prediction targets to infer AF mechanisms by comparing with ground-truth AF mechanisms that were computed using other processing methods. Ground-truth FS on LA and RA were prescribed by the simulation protocol. Ground-truth drivers on LA or RA were defined as both FS as well as reentrant sources, containing episodes with phase singularity points with a total duration of $>$640 ms (80%) out of 800 ms signal duration (Chapter 3). The ground-truth macro-reentry was selected as those with differences on dominant CLs between LA and RA less than 20 ms, with no driver on LA or RA.

In Fig 6.7, we show a comparison between the ground truth and the network prediction targets. Using the targeted CAT to tag the network prediction targets, all three classes of FS, as well as LA drivers, were correctly identified with high precision and recall, which validated the proposed approach of using passive CAT to detect LA drivers in our dataset. A majority of ground-truth macro-reentries were predicted as macro-reentries. However, in AF sustained by RA drivers and other mechanisms, an inter-atrial macro-reentrant circuit could also occur.

6.3.2 Evaluation of CAT prediction

We considered the evaluation method to assess the timing of each CAT function $CAT(t)$. As shown in Fig. 6.4(b), there could be multiple CAT events in a signal, so calculating the time difference between the CAT events was not an option. Although to evaluate the data type of an event occurrence, a timing error cost function based on dynamic time warping [224] sounds reasonable. It only provides an evaluation on the full signal, but poses difficulty in evaluating the average time error for each CAT event.

Instead, we segmented the predicted $CAT(t)$ into windows of specified lengths $20, 30, \ldots, 800$ ms, and calculated the numbers of true positive, true negative, false negative and false positive instances, which is a generic evaluation method for machine learning classifiers. For a window containing a labeled CAT, if it also contains a predicted CAT, the window was marked as true positive, and otherwise, this window was marked as a false negative. Likewise, for the window not labeled as CAT, it was marked



(a) Normalized by ground truth.

(b) Normalized by target.

(c) Normalized by ground truth.

(d) Normalized by target.

Fig. 6.7. Multi-class confusion matrices comparing the ground-truth mechanisms and the network targets that we tagged by using CATs. The scores were calculated on the concatenation of five patient test sets. FS and LA driver had high precision and recall. A high recall but low precision was associated with macro-reentry, suggesting that the alternative passive CATs also occur for AF with RA drivers. The numbers in bracket show the number of signals in each class.

as true negative if it does not contain a predicted CAT, and otherwise, a false positive. This evaluation method allowed us to inspect the model performance at different time resolutions. The probabilistic output was evaluated by the AUROC score to the probabilistic output, whereas the binary output was evaluated by accuracy, recall and precision. The evaluation on all five test datasets is shown in Fig. 6.8.

In both networks, the AUROC and accuracy scores were almost constant, irrespective of the window size, with GCRNN higher than CRNN for FS_{LA} (accuracy: 0.88 vs 0.84, AUROC: 0.94 vs 0.92), FS_{RA} (accuracy: 0.92 vs 0.86, AUROC: 0.97 vs 0.94) on a window size of 250 ms. The GCRNN also produced a smaller variation on the scores between the test patients. However, the recall and precision scores increased with larger window size, and reached their maximum value on a window size of around 250 ms,

but some of them slightly declined afterwards. The recall and precision scores were also higher in GCRNN than in CRNN, on FS_{LA} (recall: 0.85 vs 0.84, precision: 0.81 vs 0.74) and on FS_{RA} (recall: 0.92 vs 0.86, precision: 0.84 vs 0.74), evaluated on a window size of 250 ms.

For $Conn_{LA}$ and $Conn_{RA}$, the recall of $Conn_{LA}$ in GCRNN was about 5% higher than in CRNN. Little difference between these two networks was found, with AUROC and accuracy scores being about 5–10% lower than those of FS_{LA} and FS_{RA} on the window size of $\leq$250 ms. The lower AUROC also shows a 5–10% decline in the recall and precision scores in $Conn_{RA}$ on the window size of $\leq$250 ms. The reason why $Conn_{RA}$ appeared to be the worst is likely due to the drivers on LA usually having shorter CLs, which created more complex fibrillatory waves that reach the RA, and therefore, the ground-truth labels for $Conn_{RA}$ were the least organized, rending the prediction of $Conn_{RA}$ the most difficult.

6.3.3 Regular single-chamber passive CAT indicates single-chamber drivers

The passive CAT indicated the presence of active CATs that result from FSs, as shown in Fig. 6.9. In the four cases, the predicted CAT likelihoods for the FS were all small, but the presence of FSs was ascertained by the presence of regular passive CATs on the opposite atria. The four subplots show that this applies for FSs on both atria, regardless of the CL of the FS.

In Fig. 6.10 (a-b), single-chamber drivers on LA and RA were characterized by repetitive passive CATs on the opposite chamber, with CAT likelihood above the detection threshold, but no repetitive passive CAT was detected on the same chamber. We also showed two cases which successful identified macro-reentries in Fig. 6.10 (c-d), corresponds to Fig. 6.2 and Fig. 6.3, respectively, by using repetitive passive CATs in both atria.

6.3.4 Evaluation of AF mechanism prediction via CAT detection

In Fig 6.11 and Table 6.1, the presence and location of FS on both atrial walls were correctly identified by our GCRNN, with high mean scores of accuracy (89.3 – 93.3%), precision (83.9 – 88.5%) and recall (82.7 – 91.8 %) for all three classes of “FS on LA”, “FS on RA” and “No FS”, which are higher than CRNN.

In Fig 6.12 and Table 6.1, both networks had a similar performance. The presence and locations of drivers for both atrial chambers were correctly identified. Moderate high mean accuracy, precision and recall scores were seen for the two classes of “Driver

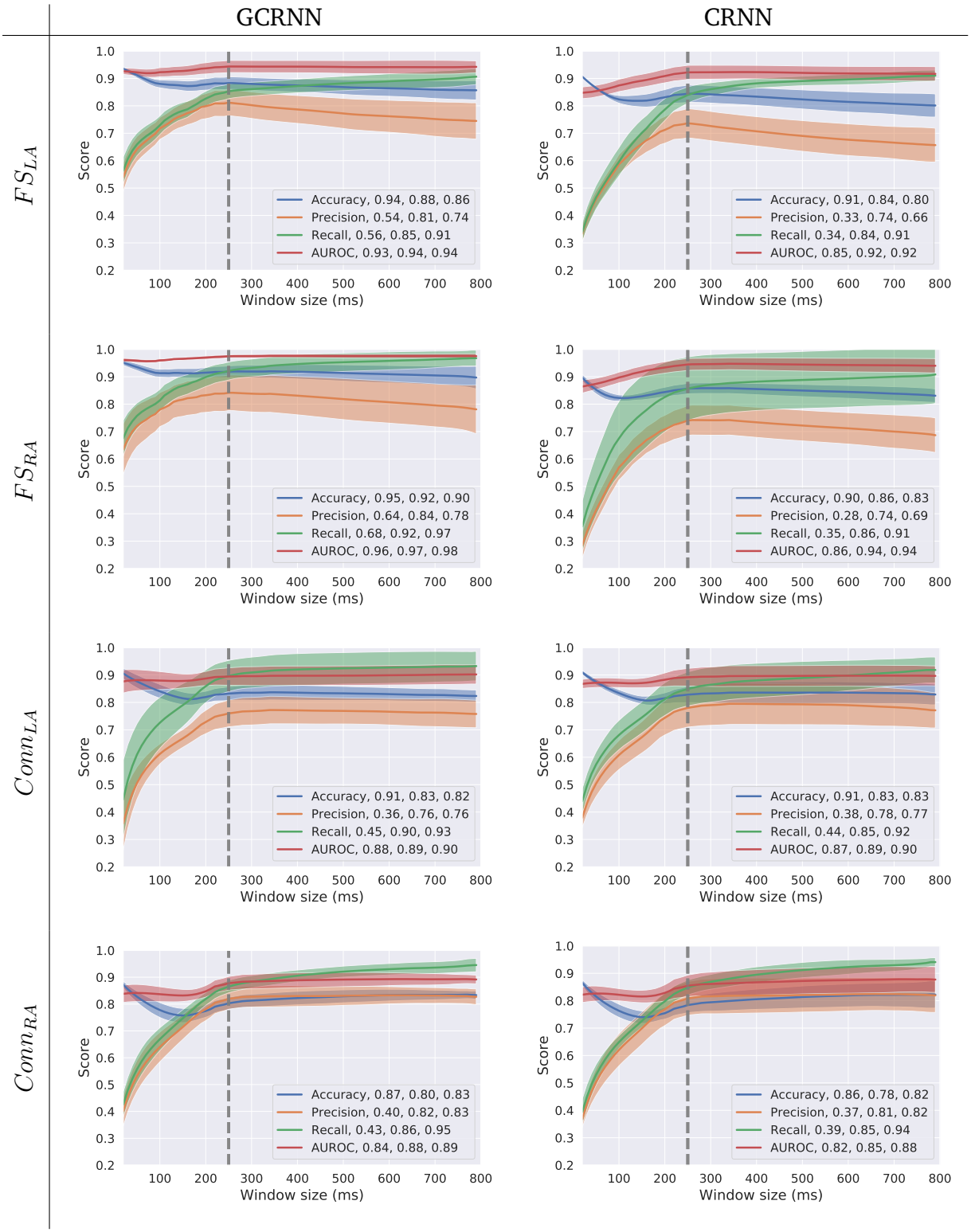
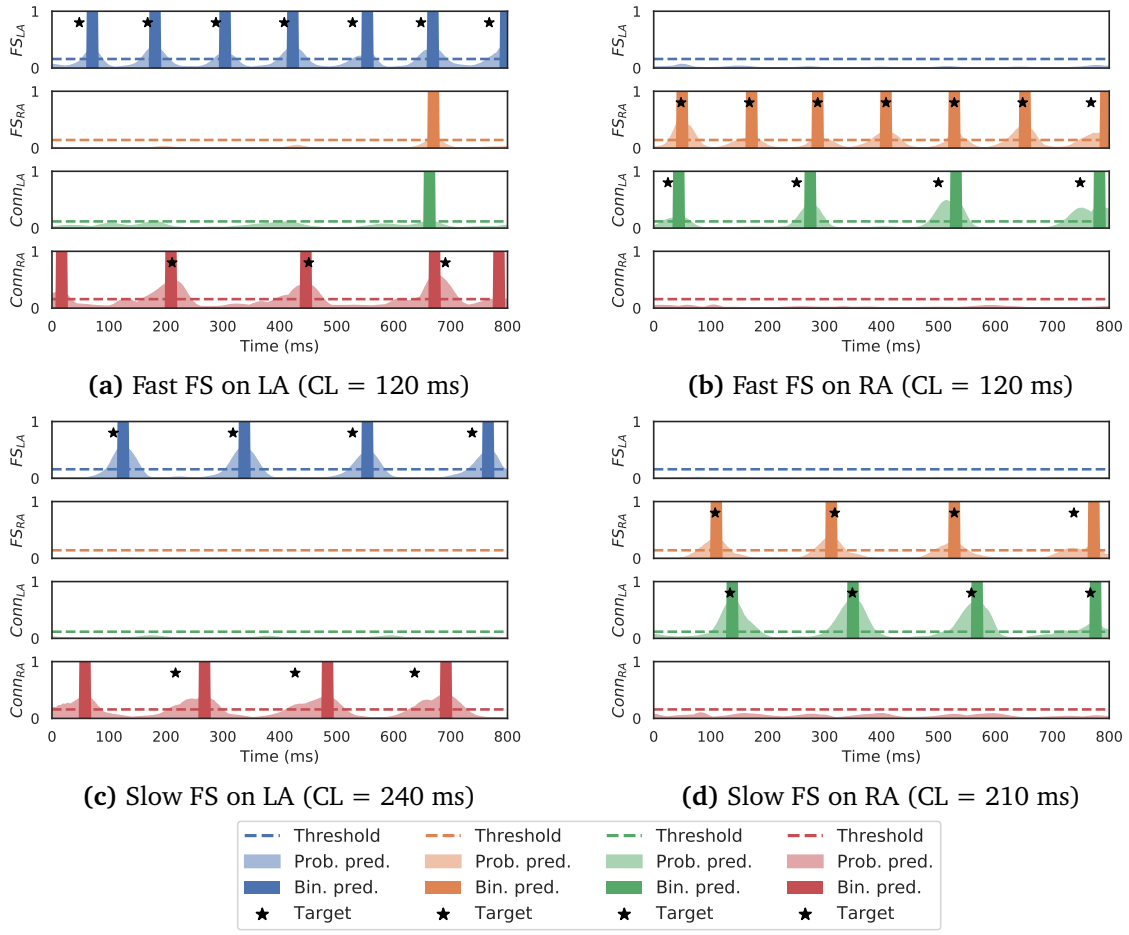


Fig. 6.8. The AUROC, accuracy, precision and recall scores on the CAT prediction outputs of (a) FS_{LA}, (b) FS_{RA}, (c) $Conn_{LA}$, and (d) $Conn_{RA}$ by GCRNN and by CRNN, on all the test sets of five patients. The shading shows the inter-patient standard deviation of the metric with the same colour as the shading. The two numerical values in the legend show the mean score of the five test patients, evaluated on windows of 20 ms, 250 ms and 800 ms, respectively.



(a) Fast FS on LA (CL = 120 ms)

(b) Fast FS on RA (CL = 120 ms)

(c) Slow FS on LA (CL = 240 ms)

(d) Slow FS on RA (CL = 210 ms)

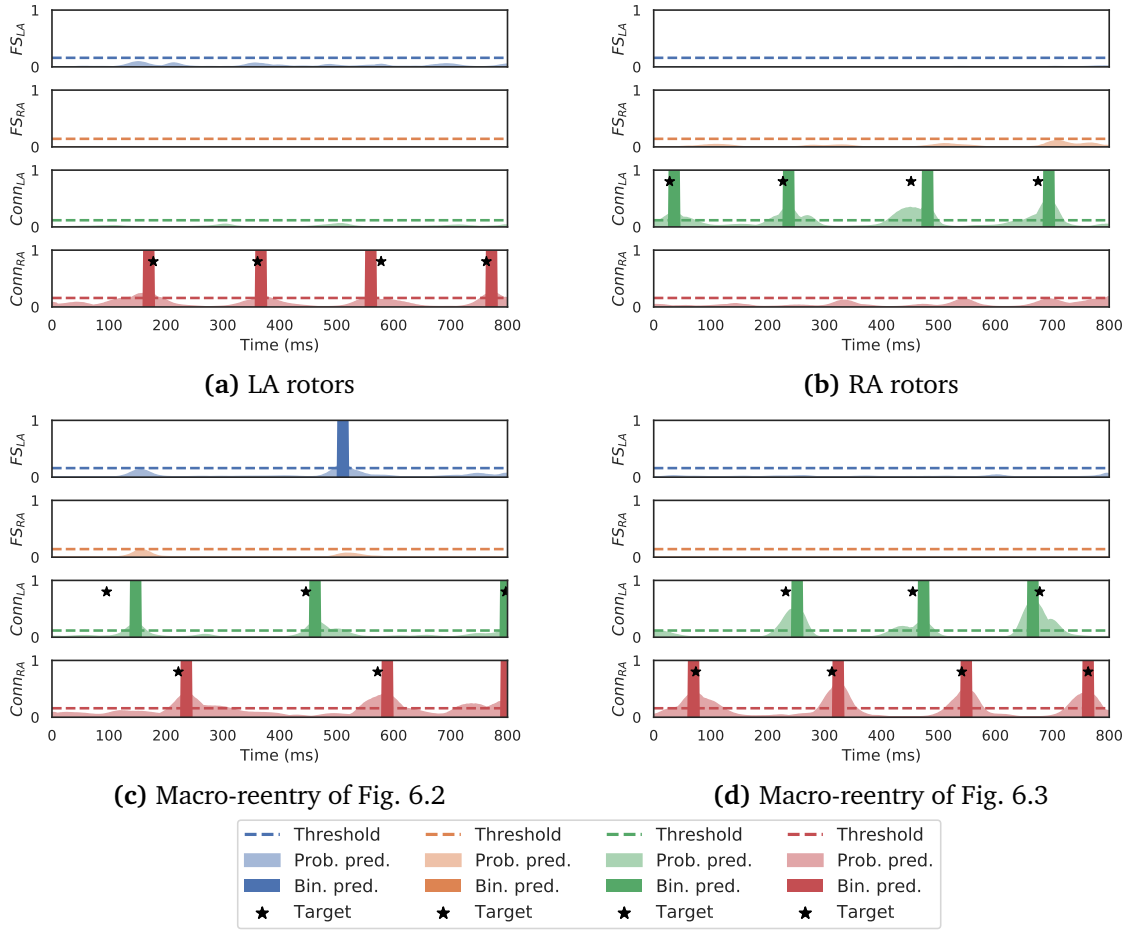
Fig. 6.9. The targeted and predicted CAT likelihoods of four FSs with (a-b) short and (c-d) long CLs by GCRNN, where the detection of passive CAT aids the detection of active CAT. The probabilistic prediction (prob. pred.) is represented as the half-transparent shaded area. The cut-off threshold is shown by the dash lines. The binary prediction (bin. pred.) is represented as the vertical bars with zero opacity. The network prediction targets are marked by black stars.

on LA" and "Driver on RA" around 70–85%. The macro-reentry received around 82% accuracy, but lower mean recall and precision scores, around 50–65%.

6.4 Discussion

We presented a deep learning model which detects not only the presence and the location, but also the onset timing of the active CAT of FS, as well as the passive CAT coming from the opposite chambers through the inter-atrial connections. The prediction output is a probabilistic function CAT(t) detailing the likelihood of CAT of each time step t.

The BSPM, as a far-field signal, is difficult to utilize for high-resolution inverse computation. The most impressive result for our method is the ability to detect the



(a) LA rotors

(b) RA rotors

(c) Macro-reentry of Fig. 6.2

(d) Macro-reentry of Fig. 6.3

Fig. 6.10. The targeted and predicted CAT likelihoods by GCRNN of four AF episodes driven by (a) rotors on LA and (b) rotors on RA, and (c-d) macro-reentries. The rotors can be detected by repetitive CAT on the opposite chamber. The legend is the same as Fig. 6.9.

Table 6.1. Accuracy, precision and recall scores (mean±standard deviation over five patient test datasets, unit:%) of FS prediction based on the active CATs of FS_{LA} and FS_{RA} as well as driver prediction based on the passive CATs of $Conn_{LA}$ and $Conn_{RA}$ by GCRNN and CRNN. The bold entries received a higher (more than 2%) mean score than the other network architecture.

	GCRNN			CRNN		
	Accuracy	**Precision**	**Recall**	**Accuracy**	**Precision**	**Recall**
FS on LA	**90.8±1.9**	**88.5±3.3**	**82.7±4.4**	87.6±2.4	83.6±3.2	77.4±6.0
FS on RA	**93.3±1.7**	**88.4±5.7**	**91.8±3.3**	86.4±1.8	76.2±5.8	85.9±11.4
No FS	**89.3±2.0**	**83.9±2.6**	**85.9±3.4**	77.5±3.0	70.9±7.5	64.4±9.5
Driver on LA	81.5±1.7	80.6±7.9	73.6±6.0	81.8±2.2	79.5±6.4	75.3±5.0
Driver on RA	84.6±2.5	71.3±10.1	**75.5±5.5**	82.8±2.4	71.6±10.4	68.6±9.8
Inter-atrial macro-reentries	82.2±3.4	49.7±5.2	**64.5±7.8**	81.4±3.4	49.6±13.5	59.9±7.8

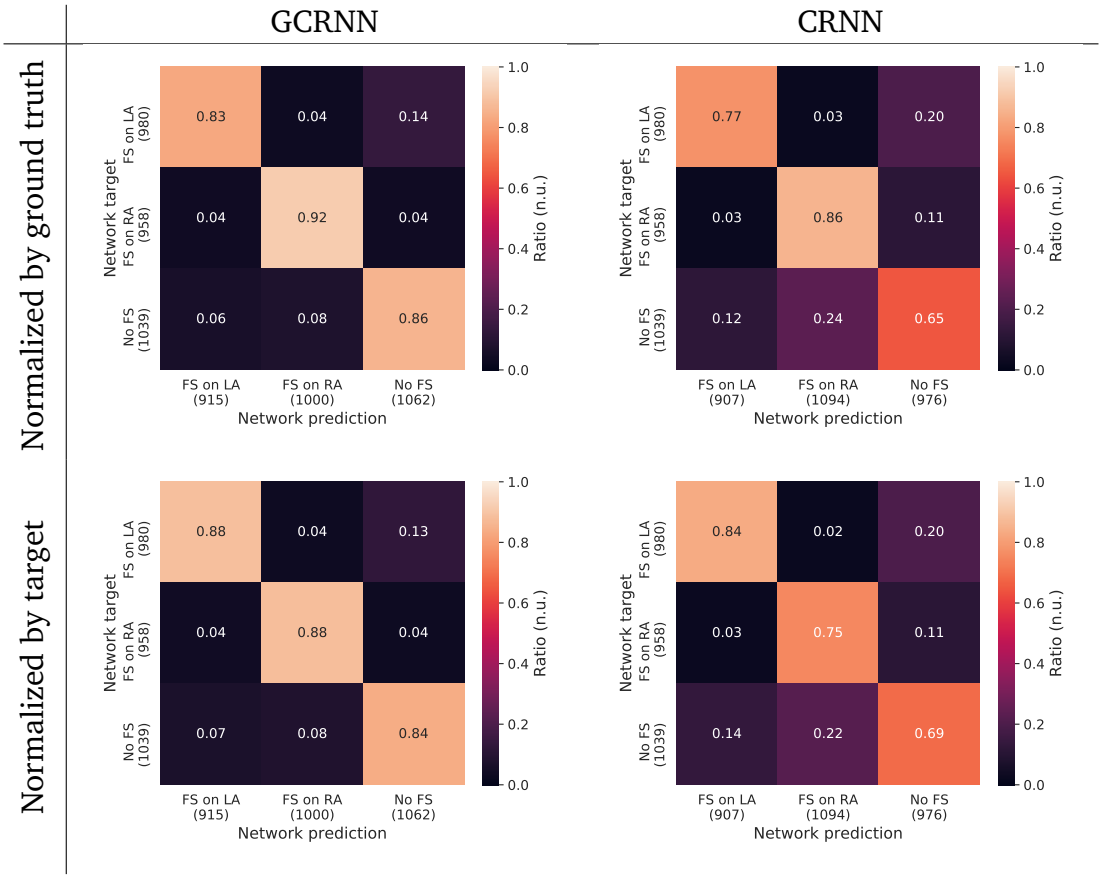


Fig. 6.11. Multi-class confusion matrices showing the performance for FS prediction based on the active CATs of FS_{LA} and FS_{RA} by GCRNN and CRNN. The scores were calculated on concatenation of five patient test sets. The numbers in bracket show the number of signals in each class.

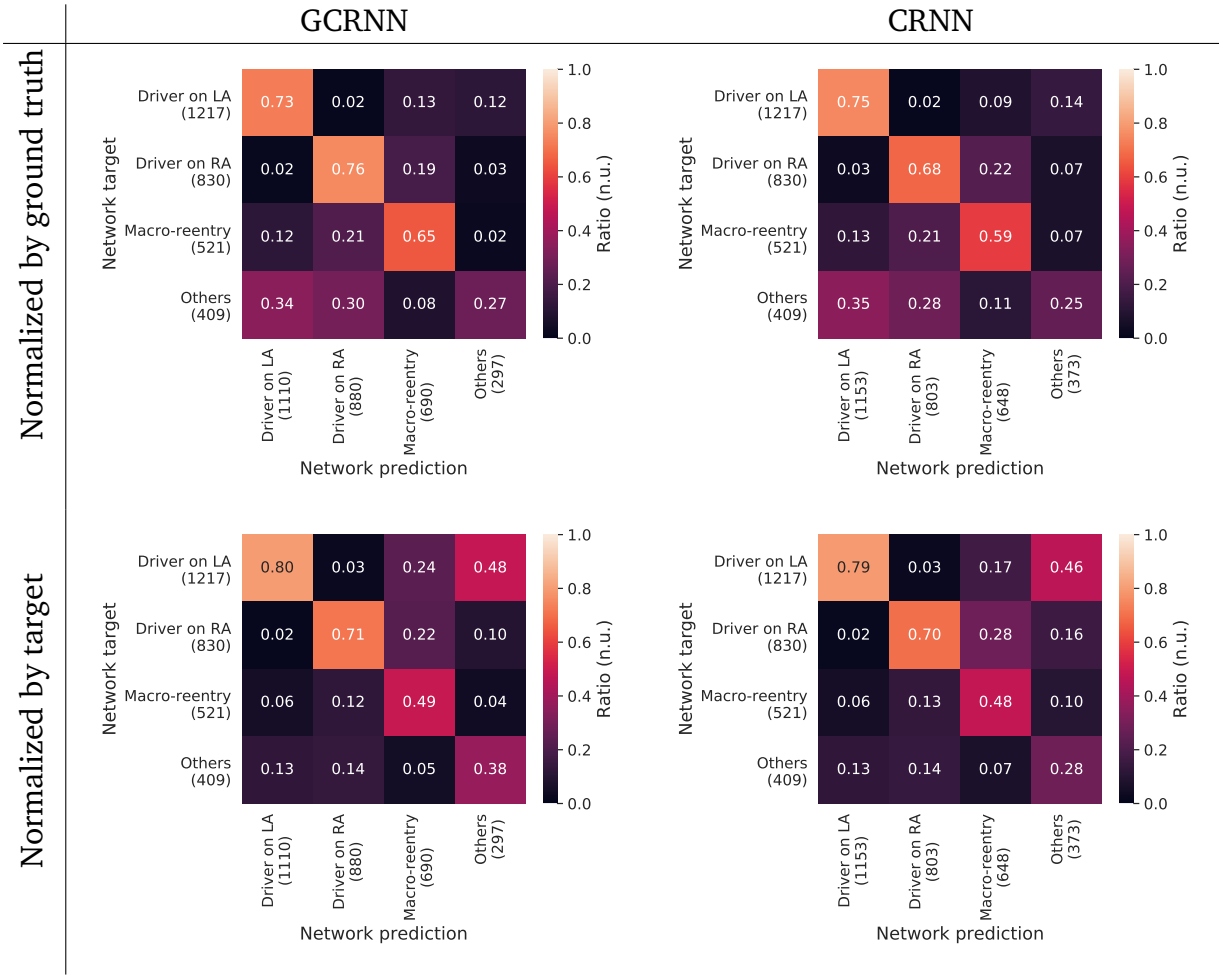


Fig. 6.12. Multi-class confusion matrices showing the performance for driver prediction based on the passive CATs of $Conn_{LA}$ and $Conn_{RA}$ by GCRNN and CRNN. The scores were calculated on concatenation of five patient test sets. The numbers in bracket show the number of signals in each class.

passive CAT of about 2 cm radius in the euclidean distance, with above 80% accuracy and almost 90% AUROC evaluated at a time window of 250 ms, which was inaccessible using ECGi. The detection of the passive CAT should be credited to the advancement in research of both geometric deep learning and high-resolution computer modeling for AF, which includes the intra-atrial connection structures into forward modeling of BSPMs.

6.4.1 Non-invasive mapping of AF mechanisms via CAT prediction

We explicitly modeled the active and passive CATs on both chambers, with four time functions predicted by our GCRNN. This way, the ectopic focal firing can be differentiated from the passive focus-like activation, with their left or right atrial location also being predicted. Combined with FS detection, the mapping time to localize these focal and reentrant drivers could be reduced. The drivers we mapped are not restricted to focal or reentrant types, and can be one or multiple drivers, but an ensemble of them on one chamber creates a repetitive dominant rhythm that passes to the other chamber via the inter-atrial connections. Moreover, not only the CL for the dominating driver can be assessed, but also its impact on the rhythm of the other atrial chambers.

We showed the validity of using passive CATs to detect single-chamber drivers, including FSs, single-chamber reentries, and macro-reentries (Fig. 6.7). All types received high precision and recall, except for macro-reentrant circuits. They can be found not only in AF with inter-atrial macro-reentry as the main driver, but also in other mechanisms such as RA drivers (Fig. 6.7 (d)). Even though we did not achieve high precision and recall for the detection of macro-reentry, our detection algorithm of the macro-reentrant circuit is of great value. The inter-atrial macro-reentry was common in FS-induced AF episodes given appropriate CL and atrial size, such as the case presented in Fig 6.3, which was induced by FS with CL of 210 ms. As there is no sustained (> 100 ms) rotor presented on the atria, it could be hard to detect such macro-reentry, since the atrial activity is regular, and as described in the introduction, ECGi did not provide enough spatial resolution to represent them.

Based on our method, the involved inter-atrial connections of the macro-reentrant circuit can also be easily inferred by extracting CLs from the predicted time-value functions of $Conn_{LA}$ and $Conn_{RA}$, as the macro-reentry travelling through the coronary sinus ostium has a longer CL than those without. An alternative method is to predict the CAT of each of the inter-atrial connections separately.

Our GCRNN can be adapted to using ECG signals as inputs with minimal changes, although the performance needs to be further tested. The distance between the ECG channel can be based on the Euclidean distance of the ECG electrodes, or by the distance in the signal space.

6.4.2 Performance of prediction of CAT timing

Evaluated by cross-validation methods, the presence and location of FSs were accurately predicted with around 90% accuracy, 80–95% recall and precision (Fig. 6.11 and Table 6.1), a significant improvement compared to using the SO-BSS method in Chapter 4. The use of EdgeConv filters in the GCRNN also significantly improved the performance for FS predictions, compared to the traditional convolutional filters in CRNN, suggesting these filters are more appropriate filters to extract the spatiotemporal information for BSPMs.

Using the $Conn_{LA}$ and $Conn_{RA}$, AF drivers dominating in a single chamber (i.e. driver on LA and driver on RA) were also successfully predicted, approaching 85% accuracy, and, 70–85% recall and precision scores (Fig. 6.12 and Table 6.1), with similar performance between GCRNN and CRNN.

We only targeted those with alternative passive CAT on both atrial chambers, which did not include other types of macro-reentries such as those circulating the tricuspid valve during an atrial flutter episode. Although we achieved only about 65% recall and 49% recall scores, we still managed to extract the inter-atrial macro-reentry circuits in some cases (Fig. 6.10), including those using Bachmann's bundle and coronary sinus as the macro-reentrant circuit (Fig. 6.2) and those using Bachmann's bundle and fossa ovalis as the macro-reentrant circuit (Fig. 6.3), which is totally inaccessible to a low-resolution mesh used by ECGi.

We also noticed that the onsets of some predicted CATs had some delays compared to the onsets of the targeted CAT, which can induce some errors when evaluating the model performance using a small time window (Fig. 6.8). This is expected as the CAT produces little change in the body surface potentials until its radius reaches a threshold, and therefore, the propagation time from the onset to this detectable stage appeared as a time lag. Nonetheless, this does not affect the prediction of AF mechanisms assessed by the number of occurrences, the CL, and the atrium that the predicted CAT functions indicate.

Similar to ECGi, some rotors were also mistaken by the neural network as the origins of an FS. This problem, however, may be solved by incorporating the amplitudes and timings of the probabilistic prediction of FS_{LA} and FS_{RA}, as well as those of passive CATs, into the mechanism prediction method. As a rotor is likely composed of different frequencies from sites on different radii, the detection of FS may be further improved by combining the prediction results with SO-BSS method introduced in Chapter 4.

6.4.3 Limitation

The way we automatically tagged CAT can introduce inaccuracy to the network labels. For example, the anisotropy in the conduction velocity due to fibre orientation was also omitted which miss some CATs. Other forms of non-centrifugal activation, such as focal breakthroughs where conduction blocks are seen in the targeted area, can also occur, both at the inter-atrial septa, and as the origins of atrial tachycardia. The detection threshold for tagging CAT could be optimized. More accurate detection of passive activation through the septum and coronary sinus could be included in future work.

The inter-patient variability of the atrial meshes in terms of location and geometry could also account for the prediction error. For further development, we could train a deep neural network from signals obtained from a larger number of atrial meshes, with smooth changes in the shapes and positions, where the impacts of geometrical and positioning on the network performance could be quantified.

Lastly, we could proceed by applying our deep learning predictor on early-stage AF patient signals. As the deep learning approach is based purely on the morphologies on the BSPM signals, it is anticipated that it could be sensitive to noise. Therefore, noises typically seen during the recording of the BSPM signal should be added to the synthetic dataset for training, to obtain a model robust to the real clinical setting.

Chapter 7

Discussion

This thesis presented three novel spatiotemporal machine learning algorithms using BSPM data with various prediction targets to improve AF treatment, without the need for patient cardiac imaging, as summarized in Fig.7.1. We specifically tackled the challenges of training and validating machine learning models on clinical datasets by leveraging synthetic AF signals generated using detailed patient atrial geometries with realistic electrophysiological properties. Furthermore, we used the synthetic datasets not only to generate training examples, but also as a deductive reasoning toolbox, to show the causal relationship for predicting the outcome of focal ablation from patient preoperative BSPMs.

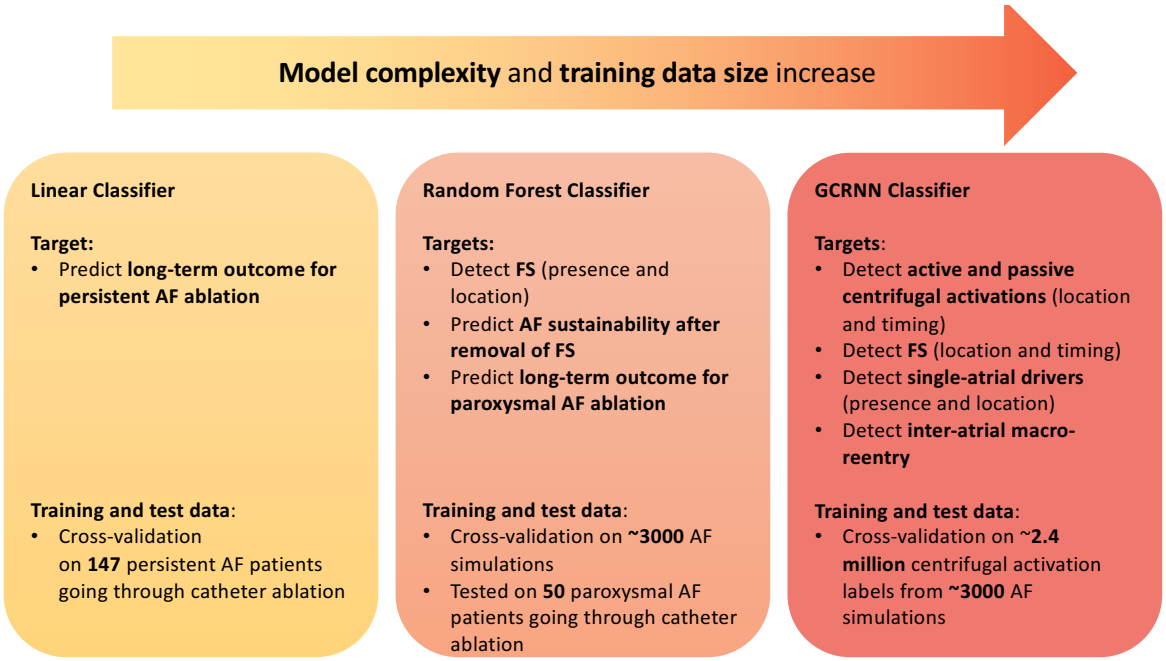


Fig. 7.1. Summary of our classifiers on BSPMs, with various degree of model complexity, and different training and testing datasets. AF: atrial fibrillation. FS: focal source.

7.1 Summary of scientific findings

We described the development of computational models of AF episodes and BSPMs in Chapter 3. This provided a total number of 2977 detailed AF simulations, covering mechanisms ranging from organized to disorganized, including episodes driven by focal source only, by reentrant sources only, by a mixture of reentrant and focal sources, as well as by multiple short-lived rotors, over five atria with differences in anatomy and location. We also showed the statistical correlation required by a machine learning model, by demonstrating the causal relationship between the pre-operative signals and prediction targets, i.e. focal ablation outcomes. We showed that 1) the rotor development suggests an arrhythmogenic substrate, represented as AF inducibility when the focal source is removed, and 2) CL of the focal source is an important indicator of its AF inducibility, with the AF-susceptible CL range differing between different focal source sites.

In Chapter 4, we used a random forest classifier, which included features of CL, MaxAC and s_1-to-signal contribution, extracted from BSPMs and ECGs using SO-BSS. We were able to determine focal and/or reentrant sources with high accuracy (above 90%), as well as the focal site with high accuracy of 80%, and demonstrate robustness to interpatient variability on the synthetic dataset in Chapter 3. Predicted from preoperative BSPMs, patients with single-atrial focal sources were associated with lower AF recurrence up to three years (p-value <0.05, logrank test) for paroxysmal AF patients following AF ablation from the Haut-Lévèque Cardiology Hospital of Bordeaux. The SO-BSS methods can be used to separate other types of signals, as well as other types of arrhythmia.

In Chapter 5, we presented a linear classifier to predict long-term AF recurrence directly using preoperative BSPMs of persistent AF patients. We developed the APSS measure as a surrogate model, which was computed from BSPMs of length $\geq$ 800 ms. APSS was developed using SO-BSS and GPR. Survival analysis was applied to select the most relevant feature from the APSS to predict the procedural outcomes. We showed that higher MaxACs, over CLs from 220–230 ms as extracted from pre-operative BSPMs, were associated with higher AF recurrence up to four years post-ablation (p-value <0.01, logrank test) for persistent AF patients from the Haut-Lévèque Hospital.

In Chapter 6, on the synthetic dataset, a deep learning model specifically designed for irregularly positioned electrodes of the mapping vest for BSPMs, GCRNN, was developed to annotate the onset timing of centrifugal activation from BSPMs. The centrifugal activations included both active activations from focal sources and passive activations through interatrial connections. The prediction output was a probabilistic function detailing the likelihood of centrifugal activation at each time step. The classifiers achieved a high accuracy (87–95%) over a time window of 20 ms. The focal source, as well as single-atrial drivers on both atria, were also predicted with high accuracy (82–94%,

respectively). Furthermore, for the first time, it was possible to non-invasively detect macro-reentries circulating through interatrial connections including Bachmann's bundle, the muscular sheath of the coronary sinus, and/or the fossa ovalis.

7.2 Impacts

7.2.1 Spatiotemporal machine learning on scanner-free imaging of body surface potentials

The most important contribution we made was the development of spatiotemporal machine learning algorithms on BSPM signals without the need for cardiac imaging. The two words "spatiotemporal" and "scanner-free" seem conflicting, but we showed that extraction of spatiotemporal structure can be done by exploiting the common structure from second-order statistics amongst the signals, or the neighbourhood information of the electrodes. This greatly simplifies the clinical procedures, as patients are no longer subject to any cardiac imaging. Additionally, it avoids the need to solve the ill-posed inversion problem of ECGi.

A surprising result in Chapter 4, was that using SO-BSS, we managed to extract spatiotemporal structure from the signals without any information of the heart-torso geometries, which can be seen in the $\mathbf{s}_1$-to-lead contributions from figures such as Fig. 4.5 and Fig. 4.7. SO-BSS was also robust to the various geometries of the heart and the torso. This suggests that BSPM signals already contain the spatiotemporal structure statistically, and thereby exempt the need to use imaging data.

The deep learning method in Chapter 6 was also spatiotemporal, as it used an EdgeConv operation which performs convolution, followed by a temporal RNN. The EdgeConv operation was based on the information of nearest electrodes, and as the relative positions of electrodes can be fixed by using a vest such as CardioInsight™ Mapping Vest, the algorithm is also free of imaging. The electrode placement may also be mapped with a general-purpose camera at a much lower cost.

7.2.2 Machine learning aided by multiscale modeling for personalized AF treatment

Another major contribution we made was the development of a paradigm in the development of machine learning algorithms on AF, which enabled both inductive reasoning of machine learning methods and deductive reasoning of multiscale modeling methods. This is especially important since the AF is a complex disease, where multiscale model-

ing was essential for personalized AF modeling. We considered the following two key elements for training prediction models using synthetic data on patients for personalized AF treatment.

First, the model should be complex enough to represent the statistical relationship between the data and the label. Meanwhile, the model complexity should also match the available data size, so that the model can be trained while avoiding overfitting. We specifically designed machine learning models that tailored the size of the training set, as shown in Fig. 7.1. From the simplest to the most complex classification models, the required number of examples also increased. The simplest model we trained and tested was a linear classifier with the covariate chosen by a Cox proportional hazard model, directly on the signals and outcomes of about 150 persistent AF patients in Chapter 5. The classifier with medium-level complexity is a nonlinear random forest classifier. It was designed to predict the different types of focal source and AF episodes seen in paroxysmal AF patients. As we only had 50 patients, we generated a synthetic data set containing nearly 3k AF episodes, and applied the classifiers on the outcome prediction for paroxysmal AF patients. The most complex model was the GCRNN model in Chapter 6. We set the timing of active and passive centrifugal activations as the prediction targets, so that each second could be used as a training example for binary prediction, so the data size was multiplied by a factor of the time length, 800 time samples in our case, which amounted to about 2.4 million examples.

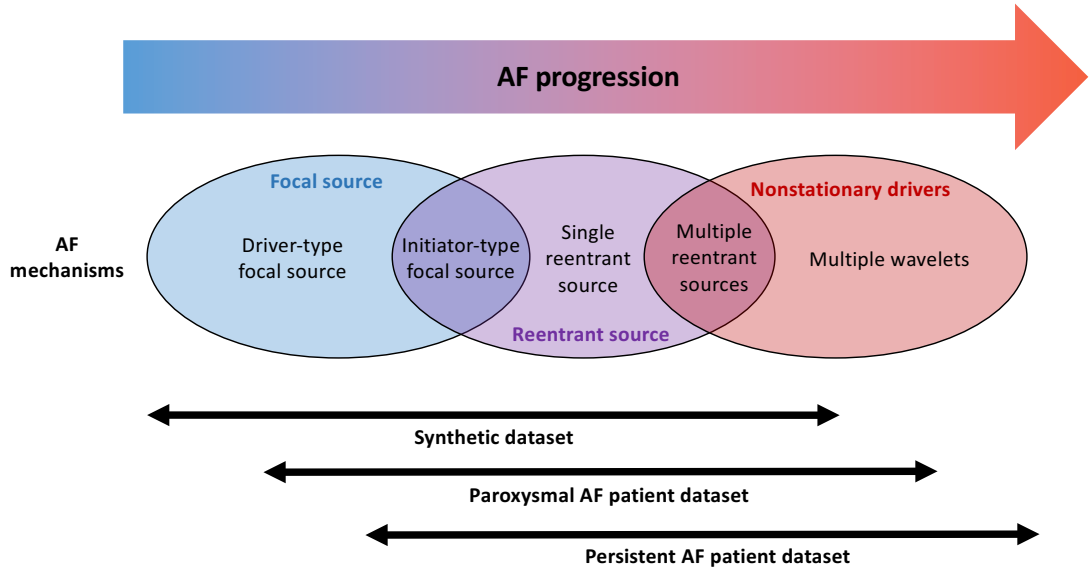


Fig. 7.2. Our synthetic datasets covered the whole spectrum of AF mechanisms listed. As we did not know the distribution of atrial activities in the paroxysmal AF and persistent AF patient datasets, we marked them at positions in the spectrum of AF progression based on our prior belief.

Second, the synthetic training data should be representative of real data, as much as possible. We specifically tackled the generality by the generation of a large number

of AF episodes with variations of the AF mechanisms, as well as atrial geometries and vest placements (Fig. 4.10). The composition of synthetic data and patient data in relation to AF progression is given in Fig. 7.2. In the synthetic data, we modeled focal sources including driver-type focal sources which drove AF through periodic ectopic activations, and initiator-type focal sources which initiated reentries where AF was self-sustained without the focal source. In addition, we also modeled reentrant sources including those containing single reentry and multiple simultaneous reentries. As the synthetic data reflects our prior belief in the distribution of paroxysmal AF patients, we trained classifiers on the synthetic data, and transferred them to the patient data. On the contrary, we directly trained and tested prediction models on the persistent AF patient data.

Finally, to enable accurate labeling of the training examples, we developed several automatic processing tools for production, as well as post-processing, of the simulations. For producing uniformly consistent locations of the focal source sites and reentrant cores, we utilized Universal Atrial Coordinate [109] to tag the focal sites and the acetylcholine sites, and phase distribution methods [186] to initiate reentries (Chapter 3). The large number of simulations also necessitated the development of fast automatic post-processing tools to quantify important dynamic properties of AF, such as the tracking of rotor activities (Chapter 3), and the extraction of passive centrifugal activation and the inter-atrial macro-reentry mechanisms (Chapter 6) from transmembrane voltage. The detection of phase singularity points and centrifugal activation are both based on tensor operations, which can be further accelerated using graphical processing units.

7.2.3 Personalized AF treatment and early AF detection

Our algorithms were tailored for personalized treatment using BSPMs and ECGs as non-invasive screening tools. For example, by predicting the focal sources and the arrhythmogenic substrate for AF maintenance from patient signals, our methods suggested personalized strategies for the foci and/or the arrhythmogenic substrate, such as disrupting the pathway of reentrant drivers, and/or limiting the available area for rotor movement (Chapter 4). We also successfully identified the presence of inter-atrial macro-reentry which suggests ablation strategies for disrupting the macro-reentrant circuit, and such suggestions were inaccessible via other non-invasive mapping methods including ECGi (Chapter 6).

Personal suggestion for mechanism-directed treatment based on pre-operative BSPMs could also be aided by APSS in Chapter 5, which is a surrogate model for describing epicardial activation dynamics of a patient on a short recording of one cardiac beat. Inference from short signals is particularly useful for describing AF, given its beat-to-beat variation. In addition, this surrogate model can be applied to other complex cardiac conditions such as ventricular arrhythmias, of which the perpetuation mechanism is

also described by a spectrum ranging from localized drivers to multiple wavelets [225].

For health management of the general population, we provided a novel mechanism-inspired method to detect the malignant ectopic beats leading to sustained AF (Chapter 4). The early detection of AF is especially valuable, as AF is the most common type of arrhythmia [1], with a declined treatment outcome with AF progression [5]. This method could also be adapted to analyze other forms of multi-lead far-field signals, such as those from a wearable device, which further increases the accessibility of our technologies.

7.3 Future work

Model validation using external datasets, and the interpretability of the model, are important steps towards clinical validation studies. Therefore, further experimental studies and retrospective data from other centers are needed. This could help to explain the electrophysiological significance of the findings, such as whether there is any specific mechanism associated with highly periodic components with CLs of 220-230 ms.

To further leverage multiscale modeling, our atrial models could be improved to include more atrial remodeling indicators, such as AF fibrosis, for a closer match of AF patients. Also, the synthetic dataset can be made to reflect the distribution in a population, by using ionic model parameters and conductivity values not as fixed values, but rather, samples from a distribution. The atrial geometries can also be interpolated using shape analysis techniques such as the one presented by Nagel et al. [226].

More sophisticated data processing tools on large meshes to enable accurate labeling of prediction targets, such as the presence of an inter-atrial macro-reentry, may also yield improvement in the machine learning models. The patient outcome prediction methods from BSPMs could be further combined with data from multiple modalities, including the non-invasive ultrasound imaging data, and patient characteristics like age and gender. Comparison between the distributions of synthetic and patient signals is also required, to improve the generality of the synthetic data. Generative models of ECG and BSPM signals of AF are also likely to achieve better prediction performance via data augmentation. Moreover, the generative model also provides privacy-preserving data sharing, which facilitates multi-center studies.

Appendices

Appendix A

Scientific publications

Peer-review journal papers

- P Langfield, **Y Feng (joint-first author)**, et al., A novel method to correct repolarization time estimation from unipolar electrograms distorted by standard filtering, *Medical Image Analysis (2021)*.

- J Corral-Acero, F Margara, M Marciniak, C Rodero, F Loncaric, **Y Feng**, et al., The 'Digital Twin' to enable the vision of prescision cardiology, *European Heart Journal (2020)*

Peer-review journal paper in review

- **Y Feng**, C Roney, J Bayer, S Niederer, M Hocini, & E Vigmond, Detection of focal source and arrhythmogenic substrate from body surface potentials to guide atrial fibrillation ablation. (Submitted in July 2021 to *PLOS Comp Biol*).

Peer-review proceeding papers

- **Y Feng**, C Roney, M Hocini, S Niederer, & E Vigmond, Robust Atrial Ectopic Beat Classification From Surface ECG Using Second-Order Blind Source Separation, *Proceeding of Computing in Cardiology (2020, Rimini)*.

- **Y Feng** & E Vigmond, Deep Multi-Label Multi-Instance Classification on 12-Lead ECG, *Proceeding of Computing in Cardiology (2020, Rimini)*.

- **Y Feng**, M Saha, M Hocini, & E Vigmond, Noninvasive One-Year Ablation Outcome Prediction for Paroxysmal Atrial Fibrillation Using Trajectories of Activation From Body Surface Potential Maps, *Proceeding of Computing in Cardiology (2019, Singapore)*.

Peer-review abstract and poster

- **Y Feng**, M Hocini, & E Vigmond, Strong Correlation between Temporal Peaks and Spatial Peaks in Body Surface Potential Maps of Paroxysmal Atrial Fibrillation Patients, *Atrial Signals Conference (2019, Bordeaux)*.

Open software

- **Y Feng**, Source code of Deep Multi-Label Multi-Instance Classification on 12-Lead ECG, *Zenodo, Proceeding of Computing in Cardiology (2020)*.

Appendix B

Neural network architecture for the ablation analysis

We replaced the graphical convolution with a 3D convolution. The input of the CNN is four-channel potential-position vector, which combines the potential in the first channel, and the three-dimensional coordinates of all the vest leads in the other three channels.

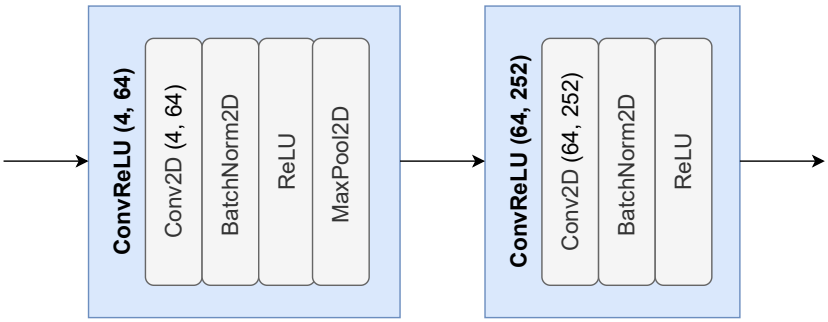


Fig. B.1. Architecture of the CNN component.

Table B.1. Hyper-parameters of the CNN component.

Component	Block	Hyper-parameters
ConvReLU (4,64)	Conv2D	In: 4, Out: 64: Kernel: (8,8), Stride: (8,8), Padding: (26,26)
	MaxPool2D	Kernel: (6,3), Stride: (6,2), Padding: (0,0)
ConvReLU (64,128)	Conv2D	In: 64, Out: 252, Kernel: (4,1), Stride: (4,1), Padding: (0,0)

The architecture and its hyperparameter are shown in Fig. B.1 and Table B.1. The CNN mainy contain *Conv2D*, *BatchNorm2D*, *ReLU* and *MaxPool2D* operations. *Conv2D* applies two-dimensional convolutional operation on the image of shape $N_{Lead} \times N_{Time}$ with a fixed-size kernel (see the "Kernel" in the Table B.1), whereas the values of this kernel are trained from the data. The *BatchNorm2D* applies normalization over a batch

so that the entries of the batch has a normal distribution of mean 0 and standard deviation of 1, rather than the whole training set, which reduces the computational cost, and empirically it was found to be a necessary component in a CNN. The *ReLU* function is a simple non-linear activation function

$$ReLU(x) = \max(0, x)$$

widely used in deep learning architectures, adding non-linearity to the network and has been found to improve classification results empirically [227]. The *MaxPool2D* module selects the maximum entry from a fixed-size region. Eventually, this CNN compressed a four-channel 252-lead 750-ms BSPM (Fig. 6.4), to a tensor of length $T_c = 100$ and input size of 252, the same output dimension as the GNN module. This tensor is then fed to the RNN module.

Bibliography

[1] Ben Freedman et al. “World Heart Federation roadmap on atrial fibrillation – A 2020 update”. In: *Global Heart* 16.1 (2021), pp. 1–36.

[2] Jason Andrade et al. “The clinical profile and pathophysiology of atrial fibrillation: Relationships among clinical features, epidemiology, and mechanisms”. In: *Circulation Research* 114.9 (2014), pp. 1453–1468.

[3] Pierre Jaïs et al. “Catheter ablation versus antiarrhythmic drugs for atrial fibrillation: The A4 study”. In: *Circulation* 118.24 (2008), pp. 2498–2505.

[4] Michel Haïssaguerre et al. “Spontaneous Initiation of Atrial Fibrillation by Ectopic Beats Originating in the Pulmonary Veins”. In: *New England Journal of Medicine* 339.10 (1998), pp. 659–666.

[5] Hugh Calkins et al. “2017 HRS/EHRA/ECAS/APHRS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation”. In: *Europace* 20.1 (2018), e1–e160.

[6] Anand N. Ganesan et al. “Long-term outcomes of catheter ablation of atrial fibrillation: a systematic review and meta-analysis.” In: *Journal of the American Heart Association* 2.2 (2013).

[7] Atul Verma et al. “Approaches to Catheter Ablation for Persistent Atrial Fibrillation”. In: *New England Journal of Medicine* 372.19 (2015), pp. 1812–1822.

[8] Pablo B. Nery et al. “Relationship Between Pulmonary Vein Reconnection and Atrial Fibrillation Recurrence: A Systematic Review and Meta-Analysis”. In: *JACC: Clinical Electrophysiology* 2.4 (2016), pp. 474–483.

[9] Fu Siong Ng et al. “Toward Mechanism-Directed Electrophenotype-Based Treatments for Atrial Fibrillation”. In: *Frontiers in Physiology* 11 (Aug. 2020).

[10] W. Einthoven. “Weiteres über das Elektrokardiogramm”. In: *Pflüger, Archiv für die Gesammte Physiologie des Menschen und der Thiere* 122.12 (May 1908), pp. 517–584.

[11] Josselin Duchateau et al. “Performance and limitations of noninvasive cardiac activation mapping”. In: *Heart Rhythm* 16.3 (2019), pp. 435–442.

[12] Maria S. Guillem et al. “Noninvasive Localization of Maximal Frequency Sites of Atrial Fibrillation by Body Surface Potential Mapping”. In: *Circulation: Arrhythmia and Electrophysiology* 6.2 (Apr. 2013), pp. 294–301.

[13] Thomas H. Everett IV et al. “Frequency domain algorithm for quantifying atrial fibrillation organization to increase defibrillation efficacy”. In: *IEEE Transactions on Biomedical Engineering* 48.9 (2001), pp. 969–978.

[14] Isabelle Nault et al. “Clinical value of fibrillatory wave amplitude on surface ECG in patients with persistent atrial fibrillation”. In: *Journal of Interventional Cardiac Electrophysiology* 26.1 (2009), pp. 11–19.

[15] Marianna Meo et al. “Noninvasive Assessment of Atrial Fibrillation Complexity in Relation to Ablation Characteristics and Outcome”. In: *Frontiers in Physiology* 9.JUL (July 2018), pp. 1–19.

[16] Christopher M. Bishop. “Pattern Recognition and Machine Learning”. In: *Journal of Electronic Imaging* 16.4 (Jan. 2007), p. 049901.

[17] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. `http://www.deeplearningbook.org`. MIT Press, 2016.

[18] Jaume Amores. “Multiple instance classification: Review, taxonomy and comparative study”. In: *Artificial Intelligence* 201 (Aug. 2013), pp. 81–105.

[19] Jorge Corral-Acero et al. *The ’Digital Twin’ to enable the vision of precision cardiology*. Mar. 2020.

[20] Fred Kusumoto. “Cardiac Anatomy and Electrophysiology”. In: *ECG Interpretation*. Cham: Springer International Publishing, 2020, pp. 3–11.

[21] Peter Spector. “Principles of cardiac electric propagation and their implications for re-entrant arrhythmias”. In: *Circulation: Arrhythmia and Electrophysiology* 6.3 (2013), pp. 655–661.

[22] Robert Plonsey and Roger C. Barr. *Bioelectricity: A quantitative approach*. 2007, pp. 1–528.

[23] A. L. Hodgkin and A. F. Huxley. “A quantitative description of membrane current and its application to conduction and excitation in nerve”. In: *The Journal of Physiology* 117.4 (Aug. 1952), pp. 500–544.

[24] Jem D. Lane and Andrew Tinker. “Have the findings from clinical risk prediction and trials any key messages for safety pharmacology?” In: *Frontiers in Physiology* 8.NOV (2017), pp. 1–11.

[25] Yung Hsin Yeh, Kristina Lemola, and Stanley Nattel. “Vagal atrial fibrillation”. In: *Acta Cardiologica Sinica* 23.1 (2007), pp. 1–12.

[26] A. Nygren et al. “Mathematical model of an adult human atrial cell: The role of K+ currents in repolarization”. In: *Circulation Research* 82.1 (1998), pp. 63–81.

[27] Marc Courtemanche, Rafael J. Ramirez, and Stanley Nattel. “Ionic mechanisms underlying human atrial action potential properties: Insights from a mathematical model”. In: *American Journal of Physiology - Heart and Circulatory Physiology* 275.1 44-1 (1998).

[28] Mary M. Maleckar et al. “Electrotonic coupling between human atrial myocytes and fibroblasts alters myocyte excitability and repolarization”. In: *Biophysical Journal* 97.8 (2009), pp. 2179–2190.

[29] Eleonora Grandi et al. “Human Atrial Action Potential and Ca 2+ Model”. In: *Circulation Research* 109.9 (Oct. 2011), pp. 1055–1066.

[30] Jussi T. Koivumäki, Topi Korhonen, and Pasi Tavi. “Impact of sarcoplasmic reticulum calcium release on calcium dynamics and action potential morphology in human atrial myocytes: A computational study”. In: *PLoS Computational Biology* 7.1 (2011).

[31] Mathias Wilhelms et al. “Benchmarking electrophysiological models of human atrial myocytes”. In: *Frontiers in Physiology* 3 JAN.November (2013).

[32] Dimitrios Filos et al. “Understanding the beat-To-beat variations of p-waves morphologies in af patients during sinus rhythm: A scoping review of the atrial simulation studies”. In: *Frontiers in Physiology* 10.JUN (2019).

[33] Bruce H. Smaill, Jichao Zhao, and Mark L. Trew. “Three-dimensional impulse propagation in myocardium: Arrhythmogenic mechanisms at the tissue level”. In: *Circulation Research* 112.5 (2013), pp. 834–848.

[34] Edward J. Vigmond, Rachel Ruckdeschel, and Natalia Trayanova. “Reentry in a morphologically realistic atrial model”. In: *Journal of Cardiovascular Electrophysiology* 12.9 (2001), pp. 1046–1054.

[35] Catalina Tobón et al. “A Three-Dimensional Human Atrial Model with Fiber Orientation. Electrograms and Arrhythmic Activation Patterns Relationship”. In: *PLoS ONE* 8.2 (2013).

[36] Ghassen Cheniti et al. “Atrial Fibrillation Mechanisms and Implications for Catheter Ablation”. In: *Frontiers in Physiology* 9.October (2018).

[37] Stanley Nattel, Brett Burstein, and Dobromir Dobrev. “Atrial remodeling and atrial fibrillation: mechanisms and implications.” In: *Circulation. Arrhythmia and electrophysiology* 1.1 (2008), pp. 62–73.

[38] Ziad F. Issa, John M. Miller, and Douglas P. Zipes. “Clinical arrhythmology and electrophysiology: A companion to braunwald’s heart disease”. In: *Clinical Arrhythmology and Electrophysiology: A Companion to Braunwald’s Heart Disease* (Jan. 2018), pp. 1–752.

[39] Pasquale Santangeli and Francis E. Marchlinski. “Techniques for the provocation, localization, and ablation of non–pulmonary vein triggers for atrial fibrillation”. In: *Heart Rhythm* 14.7 (2017), pp. 1087–1096.

[40] Phillip S Cuculich et al. “Noninvasive characterization of epicardial activation in humans with diverse atrial fibrillation patterns”. In: *Circulation* 122.14 (Oct. 2010), pp. 1364–1372.

[41] Michel Haissaguerre et al. “Driver domains in persistent atrial fibrillation.” In: *Circulation* 130.7 (Aug. 2014), pp. 530–8.

[42] Han S. Lim et al. “Complexity and Distribution of Drivers in Relation to Duration of Persistent Atrial Fibrillation”. In: *Journal of the American College of Cardiology* 69.10 (2017), pp. 1257–1269.

[43] Carola Gianni et al. “Novel concepts and approaches in ablation of atrial fibrillation: the role of non-pulmonary vein triggers.” In: *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology* 20.10 (Oct. 2018), pp. 1566–1576.

[44] Gordon K. Moe, Werner C. Rheinboldt, and J. A. Abildskov. “A computer model of atrial fibrillation”. In: *American Heart Journal* 67.2 (1964), pp. 200–220.

[45] J. L. Cox et al. “The surgical treatment of atrial fibrillation: III. Development of a definitive surgical procedure”. In: *Journal of Thoracic and Cardiovascular Surgery* 101.4 (1991), pp. 569–583.

[46] Ravi Mandapati et al. “Stable microreentrant sources as a mechanism of atrial fibrillation in the isolated sheep heart”. In: *Circulation* 101.2 (Jan. 2000), pp. 194–199.

[47] Jay Chen et al. “Dynamics of wavelets and their role in atrial fibrillation in the isolated sheep heart”. In: *Cardiovascular Research* 48.2 (2000), pp. 220–232.

[48] Charles Antzelevitch and Alexander Burashnikov. “Overview of Basic Mechanisms of Cardiac Arrhythmia”. In: *Cardiac Electrophysiology Clinics* 3.1 (2011), pp. 23–45.

[49] L. H. Frame et al. “Circus movement in the canine atrium around the tricuspid ring during experimental atrial flutter and during reentry in vitro”. In: *Circulation* 76.5 (1987), pp. 1155–1175.

[50] Hiroshi Morita et al. “Isolation of canine coronary sinus musculature from the atria by radiofrequency catheter ablation prevents induction of atrial fibrillation”. In: *Circulation: Arrhythmia and Electrophysiology* 7.6 (2014), pp. 1181–1188.

[51] Brian J Hansen et al. “Atrial fibrillation driven by micro-anatomic intramural re-entry revealed by simultaneous sub-epicardial and sub-endocardial optical mapping in explanted human hearts”. In: *European Heart Journal* 36.35 (2015), pp. 2390–2401.

[52] Omer Berenfeld and Hakan Oral. “The quest for rotors in atrial fibrillation: Different nets catch different fishes”. In: *Heart Rhythm* 9.9 (Sept. 2012), pp. 1440–1441.

[53] Philipp Sommer et al. “Successful Repeat Catheter Ablation of Recurrent Longstanding Persistent Atrial Fibrillation with Rotor Elimination as the Procedural Endpoint: A Case Series”. In: *Journal of Cardiovascular Electrophysiology* 27.3 (2016), pp. 274–280.

[54] Aleksandr Voskoboinik et al. “Revisiting pulmonary vein isolation alone for persistent atrial fibrillation: A systematic review and meta-analysis”. In: *Heart Rhythm* 14.5 (2017), pp. 661–667.

[55] Jordi Heijman et al. “Cellular and molecular electrophysiology of atrial fibrillation initiation, maintenance, and progression”. In: *Circulation Research* 114.9 (2014), pp. 1483–1499.

[56] Yasushi Akutsu et al. “Association Between Left and Right Atrial Remodeling With Atrial Fibrillation Recurrence After Pulmonary Vein Catheter Ablation in Patients With Paroxysmal Atrial Fibrillation”. In: *Circulation: Cardiovascular Imaging* 4.5 (Sept. 2011), pp. 524–531.

[57] Yalçin Gökoğlan et al. “Pulmonary Vein Antrum Isolation in Patients With Paroxysmal Atrial Fibrillation: More Than a Decade of Follow-Up”. In: *Circulation: Arrhythmia and Electrophysiology* 9.5 (May 2016).

[58] SANJIV M. NARAYAN, DAVID E. KRUMMEN, and WOUTER-JAN RAPPEL. “Clinical Mapping Approach To Diagnose Electrical Rotors and Focal Impulse Sources for Human Atrial Fibrillation”. In: *Journal of Cardiovascular Electrophysiology* 23.5 (May 2012), pp. 447–454.

[59] Michel Haïssaguerre et al. “Catheter ablation of long-lasting persistent atrial fibrillation: clinical outcome and mechanisms of subsequent arrhythmias.” In: *Journal of cardiovascular electrophysiology* 16.11 (Nov. 2005), pp. 1138–47.

[60] Michel Haïssaguerre et al. “Catheter ablation of long-lasting persistent atrial fibrillation: critical structures for termination.” In: *Journal of cardiovascular electrophysiology* 16.11 (Nov. 2005), pp. 1125–37.

[61] Michel Haïssaguerre et al. “Changes in Atrial Fibrillation Cycle Length and Inducibility During Catheter Ablation and Their Relation to Outcome”. In: *Circulation* 109.24 (June 2004), pp. 3007–3013.

[62] Felipe Atienza et al. “Real-time dominant frequency mapping and ablation of dominant frequency sites in atrial fibrillation with left-to-right frequency gradients predicts long-term maintenance of sinus rhythm”. In: *Heart Rhythm* 6.1 (2009), pp. 33–40.

[63] Felipe Atienza et al. “Comparison of radiofrequency catheter ablation of drivers and circumferential pulmonary vein isolation in atrial fibrillation: A noninferiority randomized multicenter RADAR-AF trial”. In: *Journal of the American College of Cardiology* 64.23 (2014), pp. 2455–2467.

[64] Sébastien Knecht et al. “Left atrial linear lesions are required for successful treatment of persistent atrial fibrillation”. In: *European Heart Journal* 29.19 (Oct. 2008), pp. 2359–2366.

[65] Koonlawee Nademanee et al. “A new approach for catheter ablation of atrial fibrillation: Mapping of the electrophysiologic substrate”. In: *Journal of the American College of Cardiology* 43.11 (June 2004), pp. 2044–2053.

[66] Atul Verma et al. “Selective complex fractionated atrial electrograms targeting for atrial fibrillation study (SELECT AF): a multicenter, randomized trial”. In: *Circulation. Arrhythmia and electrophysiology* 7.1 (Feb. 2014), pp. 55–62.

[67] Manav Sohal et al. “Is Mapping of Complex Fractionated Electrograms Obsolete?” In: *Arrhythmia & Electrophysiology Review* 2.2 (2015), pp. 109–15.

[68] Iain Sim et al. “Left atrial voltage mapping: defining and targeting the atrial fibrillation substrate”. In: *Journal of Interventional Cardiac Electrophysiology* 56.3 (Dec. 2019), pp. 213–227.

[69] Jaakko Malmivuo and Robert Plonsey. “Bioelectromagnetism Principles and Applications of Bioelectric”. In: *Oxford University Press* (1995), p. 512.

[70] Jorge G. Quintanilla et al. “Mechanistic approaches to detect, target, and ablate the drivers of atrial fibrillation”. In: *Circulation: Arrhythmia and Electrophysiology* 9.1 (2016), pp. 1–11.

[71] CE. Rasmussen and CKI. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. Cambridge, MA, USA: MIT Press, Jan. 2006, p. 248.

[72] Yingjing Feng et al. “An efficient cardiac mapping strategy for radiofrequency catheter ablation with active learning”. In: *International Journal of Computer Assisted Radiology and Surgery* 12.7 (July 2017), pp. 1199–1207.

[73] Sam Coveney et al. “Probabilistic Interpolation of Uncertain Local Activation Times on Human Atrial Manifolds”. In: *IEEE Transactions on Biomedical Engineering* (2019), pp. 1–1.

[74] Yuyang Wang, Roni Khardon, and Pavlos Protopapas. “Nonparametric Bayesian estimation of periodic light curves”. In: *Astrophysical Journal* 756.1 (2012).

[75] Nicolas Durrande et al. “Detecting periodicities with gaussian processes”. In: *PeerJ Computer Science* 2016.4 (2016), pp. 1–18.

[76] Felipe Tobar. “Bayesian nonparametric spectral estimation”. In: *Advances in Neural Information Processing Systems* 2018-Decem.NeurIPS (2018), pp. 10127–10137.

[77] Christopher J James and Christian W Hesse. *Independent component analysis for biomedical signals*. 2005.

[78] Lang Tong et al. “Indeterminacy and identifiability of blind identification”. In: *IEEE transactions on circuits and systems* 38.5 (1991), pp. 499–509.

[79] Seungjin Choi et al. “Blind source separation and independent component analysis: A review”. In: *Neural Information Processing-Letters and Reviews* 6.1 (2005), pp. 1–57.

[80] L. Tong et al. “AMUSE: A new blind identification algorithm”. In: *Proceedings - IEEE International Symposium on Circuits and Systems* 3 (1990), pp. 1784–1787.

[81] Adel Belouchrani et al. “A blind source separation technique using second-order statistics”. In: *IEEE Transactions on Signal Processing* 45.2 (1997), pp. 434–444.

[82] Aapo Hyvärinen and Erkki Oja. “A Fast Fixed-Point Algorithm for Independent Component Analysis”. In: *Neural Computation* 9.7 (1997), pp. 1483–1492.

[83] Aapo Hyvärinen. “Fast and robust fixed-point algorithms for independent component analysis”. In: *IEEE Transactions on Neural Networks* 10.3 (1999), pp. 626–634.

[84] A. J. Bell and T. J. Sejnowski. “An information-maximization approach to blind separation and blind deconvolution.” In: *Neural computation* 7.6 (1995), pp. 1129–1159.

[85] A. Hyvärinen and E. Oja. “Independent component analysis: algorithms and applications”. In: *Neural Networks* 13.4-5 (June 2000), pp. 411–430.

[86] Leo Breiman. “Random Forests”. In: *Machine Learning* 45.1 (2001), pp. 5–32.

[87] David E Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by back-propagating errors”. In: *Nature* 323.6088 (Oct. 1986), pp. 533–536.

[88] Kevin Jarrett et al. “What is the best multi-stage architecture for object recognition?” In: *Proceedings of the IEEE International Conference on Computer Vision* (2009), pp. 2146–2153.

[89] Sergey Ioffe and Christian Szegedy. “Batch normalization: Accelerating deep network training by reducing internal covariate shift”. In: *32nd International Conference on Machine Learning, ICML 2015*. Vol. 1. 2015, pp. 448–456.

[90] Zhou and Chellappa. “Computation of optical flow using a neural network”. In: *IEEE International Conference on Neural Networks*. 86. IEEE, 1988, pp. 71–78.

[91] Kyunghyun Cho et al. “Learning phrase representations using RNN encoder-decoder for statistical machine translation”. In: *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference* (2014), pp. 1724–1734.

[92] Sepp Hochreiter and Jürgen Schmidhuber. “Long Short-Term Memory”. In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780.

[93] Yulan Guo et al. “Deep Learning for 3D Point Clouds: A Survey”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020), pp. 1–1.

[94] Charles R. Qi et al. “PointNet: Deep learning on point sets for 3D classification and segmentation”. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-Janua (2017), pp. 77–85.

[95] Yongcheng Liu et al. “Relation-shape convolutional neural network for point cloud analysis”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2019-June (2019), pp. 8887–8896.

[96] Jonathan Masci et al. “Geometric deep learning”. In: *SIGGRAPH ASIA 2016 Courses*. July. New York, NY, USA: ACM, Nov. 2016, pp. 1–50.

[97] Justin Gilmer et al. “Neural message passing for quantum chemistry”. In: *34th International Conference on Machine Learning, ICML 2017* 3 (2017), pp. 2053–2070.

[98] Matthias Fey and Jan Eric Lenssen. “Fast Graph Representation Learning with PyTorch Geometric”. In: 1 (2019), pp. 1–9.

[99] Yue Wang et al. “Dynamic graph CNN for learning on point clouds”. In: *arXiv* 1.1 (2018).

[100] D. Harrild and C. Henriquez. “A computer model of normal conduction in the human atria.” In: *Circulation research* 87.7 (2000).

[101] Edward J. Vigmond et al. “The effect of vagally induced dispersion of action potential duration on atrial arrhythmogenesis”. In: *Heart Rhythm* 1.3 (2004), pp. 334–344.

[102] Vincent Jacquemet et al. “Study of Unipolar Electrogram Morphology in a Computer Model of Atrial Fibrillation”. In: *Journal of Cardiovascular Electrophysiology* 14.10 SUPPL. (2003), pp. 172–179.

[103] Cesare Corrado et al. “A work flow to build and validate patient specific left atrium electrophysiology models from catheter measurements”. In: *Medical Image Analysis* 47 (2018), pp. 153–163.

[104] Martin W. Krueger et al. “Personalization of atrial anatomy and electrophysiology as a basis for clinical modeling of radio-frequency ablation of atrial fibrillation”. In: *IEEE Transactions on Medical Imaging* 32.1 (2013), pp. 73–84.

[105] Thomas E. Fastl et al. “Personalized computational modeling of left atrial geometry and transmural myofiber architecture”. In: *Medical Image Analysis* 47 (2018), pp. 180–190.

[106] Simon Labarthe et al. “A bilayermodel of human atria:mathematical background, construction, and assessment”. In: *Europace* 16.May (2014), pp. iv21–iv29.

[107] Caroline H Roney et al. “Variability in pulmonary vein electrophysiology and fibrosis determines arrhythmia susceptibility and dynamics”. In: *PLOS Computational Biology* 14.5 (May 2018). Ed. by Alison L Marsden, e1006166.

[108] Kathleen S. McDowell et al. “Virtual electrophysiological study of atrial fibrillation in fibrotic remodeling”. In: *PLoS ONE* 10.2 (2015), pp. 1–16.

[109] Caroline H. Roney et al. “Universal atrial coordinates applied to visualisation, registration and construction of patient specific meshes”. In: *Medical Image Analysis* 55 (July 2019), pp. 65–75.

[110] Konstantinos N. Aronis, Rheeda Ali, and Natalia A. Trayanova. “The role of personalized atrial modeling in understanding atrial fibrillation mechanisms and improving treatment”. In: *International Journal of Cardiology* 287 (July 2019), pp. 139–147.

[111] Sohail Zahid et al. “Patient-derived models link re-entrant driver localization in atrial fibrillation to fibrosis spatial pattern”. In: *Cardiovascular Research* 110.3 (2016), pp. 443–454.

[112] Caroline H Roney et al. “Modelling methodology of atrial fibrosis affects rotor dynamics and electrograms”. In: *EP Europace* 18.suppl_4 (Dec. 2016), pp. iv146–iv155.

[113] Vincent Jacquemet, Nathalie Virag, and Lukas Kappenberger. “Wavelength and vulnerability to atrial fibrillation: Insights from a computer model of human atria”. In: *Europace* 7.SUPPL. 2 (2005).

[114] Michel Haissaguerre et al. “Atrial fibrillatory cycle length: computer simulation and potential clinical importance.” In: *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology* 9 Suppl 6 (2007), pp. 64–70.

[115] Yunfan Gong et al. “Mechanism underlying initiation of paroxysmal atrial flutter/atrial fibrillation by ectopic foci: A simulation study”. In: *Circulation* 115.16 (2007), pp. 2094–2102.

[116] Mirabeau Saha et al. “Wavelength and fibrosis affect phase singularity locations during atrial fibrillation”. In: *Frontiers in Physiology* 9.SEP (2018), pp. 1–12.

[117] Jason D. Bayer et al. “Novel radiofrequency ablation strategies for terminating atrial fibrillation in the left atrium: A simulation study”. In: *Frontiers in Physiology* 7.APR (2016), pp. 1–13.

[118] Patrick M. Boyle et al. “Computationally guided personalized targeted ablation of persistent atrial fibrillation”. In: *Nature Biomedical Engineering* 3.11 (2019), pp. 870–879.

[119] Axel Loewe et al. “Patient-specific identification of atrial flutter vulnerability-a computational approach to reveal latent reentry pathways”. In: *Frontiers in Physiology* 10.JAN (2019), pp. 1–15.

[120] Julie K. Shade et al. “Preprocedure Application of Machine Learning and Mechanistic Simulations Predicts Likelihood of Paroxysmal Atrial Fibrillation Recurrence following Pulmonary Vein Isolation”. In: *Circulation: Arrhythmia and Electrophysiology* 13.7 (2020), pp. 617–627.

[121] Raúl Alcaraz and José Joaquín Rieta. “Non-invasive organization variation assessment in the onset and termination of paroxysmal atrial fibrillation”. In: *Computer Methods and Programs in Biomedicine* 93.2 (2009), pp. 148–154.

[122] Antonio Hernández et al. “Preoperative study of the surface ECG for the prognosis of atrial fibrillation maze surgery outcome at discharge”. In: *Physiological Measurement* 35.7 (2014), pp. 1409–1423.

[123] Marianna Meo et al. “Spatial variability of the 12-lead surface ECG as a tool for noninvasive prediction of catheter ablation outcome in persistent atrial fibrillation”. In: *IEEE Transactions on Biomedical Engineering* 60.1 (2013), pp. 20–27.

[124] Giorgio Luongo et al. “Non-Invasive Characterization of Atrial Flutter Mechanisms Using Recurrence Quantification Analysis on the ECG: A Computational Study”. In: *IEEE Transactions on Biomedical Engineering* 68.3 (2021), pp. 914–925.

[125] Federica Censi et al. “P-wave Variability and Atrial Fibrillation”. In: *Scientific Reports* 6 (2016), pp. 1–7.

[126] Danila Potyagaylo et al. “ECG imaging of focal atrial excitation: Evaluation in a realistic simulation setup”. In: *Computing in Cardiology* 43 (2016), pp. 113–116.

[127] Ana Ferrer-Albero et al. “Non-invasive localization of atrial ectopic beats by using simulated body surface P-wave integral maps”. In: *PLoS ONE* 12.7 (2017).

[128] Eduardo Jorge Godoy et al. “Atrial fibrosis hampers non-invasive localization of atrial ectopic foci from multi-electrode signals: A 3D Simulation Study”. In: *Frontiers in Physiology* 9.MAY (2018), p. 404.

[129] Vincent Jacquemet et al. “The equivalent dipole used to characterize atrial fibrillation”. In: *Computers in Cardiology*. Vol. 33. 2006, pp. 149–152.

[130] Erick A. Perez Alday et al. “A New Algorithm to Diagnose Atrial Ectopic Origin from Multi Lead ECG Systems - Insights from 3D Virtual Human Atria and Torso”. In: *PLOS Computational Biology* 11.1 (Jan. 2015). Ed. by Alexander V. Panfilov, e1004026.

[131] Erick A. Perez Alday et al. “Novel non-invasive algorithm to identify the origins of re-entry and ectopic foci in the atria from 64-lead ECGs: A computational study”. In: *PLOS Computational Biology* 13.3 (Mar. 2017). Ed. by Alexander V Panfilov, e1005270.

[132] Daniele Giacopelli et al. “Spatial pattern of P waves in paroxysmal atrial fibrillation patients in sinus rhythm and controls”. In: *PACE - Pacing and Clinical Electrophysiology* 35.7 (2012), pp. 819–826.

[133] Jason Ng et al. “Surface ECG vector characteristics of organized and disorganized atrial activity during atrial fibrillation”. In: *Journal of Electrocardiology* 37.SUPPL. (2004), pp. 91–97.

[134] Samuel Ruipérez-Campillo et al. “Non-invasive characterisation of macroreentrant atrial tachycardia types from a vectorcardiographic approach with the slow conduction region as a cornerstone”. In: *Computer Methods and Programs in Biomedicine* 200 (2021).

[135] Miguel Ángel Cámara-Vázquez et al. “Non-invasive Estimation of Atrial Fibrillation Driver Position With Convolutional Neural Networks and Body Surface Potentials”. In: *Frontiers in Physiology* 12.October (2021), pp. 1–11.

[136] Frederique J Vanheusden et al. “Systematic differences of non-invasive dominant frequency estimation compared to invasive dominant frequency estimation in atrial fibrillation”. In: *Computers in Biology and Medicine* 104 (2019), pp. 299–309.

[137] Anne Vakkuri et al. “Time-frequency balanced spectral entropy as a measure of anesthetic drug effect in central nervous system during sevoflurane, propofol, and thiopental anesthesia”. In: *Acta Anaesthesiologica Scandinavica* 48.2 (2004), pp. 145–153.

[138] A Liberos et al. “Phase singularity point tracking for the identification of typical and atypical flutter patients: A clinical-computational study”. In: *Computers in Biology and Medicine* 104 (2019), pp. 319–328.

[139] Victor Gonçalves Marques et al. “Characterization of atrial arrhythmias in body surface potential mapping: A computational study”. In: *Computers in Biology and Medicine* 127 (2020), p. 103904.

[140] J.J. Rieta et al. “Atrial Activity Extraction for Atrial Fibrillation Analysis Using Blind Source Separation”. In: *IEEE Transactions on Biomedical Engineering* 51.7 (July 2004), pp. 1176–1186.

[141] F. Castells et al. “Spatiotemporal Blind Source Separation Approach to Atrial Activity Estimation in Atrial Tachyarrhythmias”. In: *IEEE Transactions on Biomedical Engineering* 52.2 (Feb. 2005), pp. 258–267.

[142] Raul Llinares et al. “Independent component analysis of body surface potential mapping recordings with atrial fibrillation”. In: *IEEE International Conference on Neural Networks - Conference Proceedings* (2006), pp. 2287–2294.

[143] Raul Llinares and Jorge Igual. “Exploiting periodicity to extract the atrial activity in atrial arrhythmias”. In: *Eurasip Journal on Advances in Signal Processing* 2011.1 (2011).

[144] Joshua S. Richman and J. Randall Moorman. “Physiological time-series analysis using approximate entropy and sample entropy”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 278.6 (2000), H2039–H2049.

[145] Bin He and Richard J. Cohen. “Body Surface Laplacian ECG Mapping”. In: *IEEE Transactions on Biomedical Engineering* 39.11 (1992), pp. 1179–1191.

[146] J. Lian et al. “Body surface Laplacian mapping of atrial depolarization in healthy human subjects”. In: *Medical and Biological Engineering and Computing* 40.6 (2002), pp. 650–659.

[147] Zachi I. Attia et al. “An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction”. In: *The Lancet* 394.10201 (2019), pp. 861–867.

[148] Thomas H. Everett IV et al. “Assessment of global atrial fibrillation organization to optimize timing of atrial defibrillation”. In: *Circulation* 103.23 (2001), pp. 2857–2861.

[149] Miguel Rodrigo et al. “Technical considerations on phase mapping for identification of atrial reentrant activity in direct-And inverse-computed electrograms”. In: *Circulation: Arrhythmia and Electrophysiology* 10.9 (2017).

[150] Dong Wook Kim et al. “Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: Results from recently published papers”. In: *Korean Journal of Radiology* 20.3 (2019), pp. 405–410.

[151] Olaf Dössel et al. “Computer Modeling of the Heart for ECG Interpretation—A Review”. In: *Hearts* 2.3 (2021), pp. 350–368.

[152] Axel Loewe et al. “Influence of the earliest right atrial activation site and its proximity to interatrial connections on P-wave morphology”. In: *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology* 18 (2016), pp. iv35–iv43.

[153] Robin Andlauer et al. “Influence of left atrial size on P-wave morphology: Differential effects of dilation and hypertrophy”. In: *Europace* 20 (2018), pp. III36–III44.

[154] Brian J. Hansen et al. “Maintenance of Atrial Fibrillation: Are Reentrant Drivers with Spatial Stability the Key?” In: *Circulation: Arrhythmia and Electrophysiology* 9.10 (Oct. 2016).

[155] Sigfus Gizurarson et al. “Hierarchical Schema for Identifying Focal Electrical Sources During Human Atrial Fibrillation: Implications for Catheter-Based Atrial Substrate Ablation”. In: *JACC: Clinical Electrophysiology* 2.6 (2016), pp. 656–666.

[156] Isabelle C. Van Gelder and Martin E.W. Hemels. “The progressive nature of atrial fibrillation: A rationale for early restoration and maintenance of sinus rhythm”. In: *Europace* 8.11 (2006), pp. 943–949.

[157] C F Tsai et al. “Initiation of atrial fibrillation by ectopic beats originating from the superior vena cava: electrophysiological characteristics and results of radiofrequency ablation.” In: *Circulation* 102.1 (July 2000), pp. 67–74.

[158] V A Folarin, P J Fitzsimmons, and W B Kruyer. “Holter monitor findings in asymptomatic male military aviators without structural heart disease.” In: *Aviation, space, and environmental medicine* 72.9 (Sept. 2001), pp. 836–8.

[159] Nobuyuki Murakoshi et al. “Prognostic impact of supraventricular premature complexes in community-based health checkups: The Ibaraki Prefectural Health Study”. In: *European Heart Journal* 36.3 (2015), pp. 170–178.

[160] Zeynep Binici et al. “Excessive supraventricular ectopic activity and increased risk of atrial fibrillation and stroke”. In: *Circulation* 121.17 (May 2010), pp. 1904–1911.

[161] Shinya Suzuki et al. “Usefulness of frequent supraventricular extrasystoles and a high CHADS 2 score to predict first-time appearance of atrial fibrillation”. In: *American Journal of Cardiology* 111.11 (2013), pp. 1602–1607.

[162] Thomas A. Dewland et al. “Atrial Ectopy as a Predictor of Incident Atrial Fibrillation”. In: *Annals of Internal Medicine* 159.11 (Dec. 2013), p. 721.

[163] Roy M. John and William G. Stevenson. “Predicting atrial fibrillation: can we shape the future?” In: *European Heart Journal* 36.3 (Jan. 2015), pp. 145–147.

[164] Hakan Oral et al. “Pulmonary vein isolation for paroxysmal and persistent atrial fibrillation”. In: *Circulation* 105.9 (Mar. 2002), pp. 1077–1081.

[165] Caroline H. Roney et al. “Patient-specific simulations predict efficacy of ablation of interatrial connections for treatment of persistent atrial fibrillation”. In: *EP Europace* 20.suppl_3 (Nov. 2018), pp. iii55–iii68.

[166] Siew Yen Ho et al. “Anatomy of the left atrium: Implications for radiofrequency ablation of atrial fibrillation”. In: *Journal of Cardiovascular Electrophysiology* 10.11 (1999), pp. 1525–1533.

[167] S Y Ho. “Architecture of the pulmonary veins: relevance to radiofrequency ablation”. In: *Heart* 86.3 (Sept. 2001), pp. 265–270.

[168] Mark Potse et al. “A comparison of monodomain and bidomain reaction-diffusion models for action potential propagation in the human heart”. In: *IEEE Transactions on Biomedical Engineering* 53.12 (2006), pp. 2425–2435.

[169] Edward J Vigmond et al. “Computational tools for modeling electrical activity in cardiac tissue”. In: *Journal of Electrocardiology* 36.SUPPL. (Dec. 2003), pp. 69–74.

[170] Jason D. Bayer et al. “Acetylcholine Delays Atrial Activation to Facilitate Atrial Fibrillation”. In: *Frontiers in Physiology* 10 (Sept. 2019), p. 1105.

[171] Martin J Bishop and Gernot Plank. “Bidomain ECG Simulations Using an Augmented Monodomain Model for the Cardiac Source”. In: *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING* 58.8 (2011).

[172] Frank M. Weber et al. “Predicting tissue conductivity influences on body surface potentials-An efficient approach based on principal component analysis”. In: *IEEE Transactions on Biomedical Engineering* 58.2 (2011), pp. 265–273.

[173] Charulatha Ramanathan and Yoram Rudy. “Electrocardiographic imaging: II. Effect of torso inhomogeneities on noninvasive reconstruction of epicardial potentials, electrograms, and isochrones”. In: *Journal of Cardiovascular Electrophysiology* 12.2 (2001), pp. 241–252.

[174] Satish Balay et al. “Efficient Management of Parallelism in Object Oriented Numerical Software Libraries”. In: *Modern Software Tools in Scientific Computing*. Ed. by E. Arge, A. M. Bruaset, and H. P. Langtangen. Birkhauser Press, 1997, pp. 163–202.

[175] Satish Balay et al. *PETSc/TAO Users Manual*. Tech. rep. ANL-21/39 - Revision 3.16. Argonne National Laboratory, 2021.

[176] Satish Balay et al. *PETSc Web page*. https://petsc.org/. 2021.

[177] Richard L Graham, Timothy S Woodall, and Jeffrey M Squyres. “Open MPI: A Flexible High Performance MPI”. In: *Parallel Processing and Applied Mathematics*. Ed. by Roman Wyrzykowski et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 228–239.

[178] David E. Krummen et al. “Mechanisms of human atrial fibrillation initiation clinical and computational studies of repolarization restitution and activation latency”. In: *Circulation: Arrhythmia and Electrophysiology* 5.6 (2012), pp. 1149–1159.

[179] Gunnar Seemann et al. “Heterogeneous three-dimensional anatomical and electrophysiological model of human atria”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 364.1843 (2006), pp. 1465–1481.

[180] Oleg V. Aslanidi et al. “3D virtual human atria: A computational platform for studying clinical atrial fibrillation”. In: *Progress in Biophysics and Molecular Biology* 107.1 (2011), pp. 156–168.

[181] A. Dorn et al. “Modelling of heterogeneous human atrial electrophysiology”. In: *Biomedizinische Technik* 57.SUPPL. 1 TRACK-O (2012), pp. 350–353.

[182] Martin W. Krueger et al. “In-silico modeling of atrial repolarization in normal and atrial fibrillation remodeled state”. In: *Medical and Biological Engineering and Computing* 51.10 (2013), pp. 1105–1119.

[183] Vadim V. Fedorov et al. “Complex interactions between the sinoatrial node and atrium during reentrant arrhythmias in the canine heart”. In: *Circulation* 122.8 (Aug. 2010), pp. 782–789.

[184] James Kneller et al. “Cholinergic atrial fibrillation in a computer model of a two-dimensional sheet of canine atrial cells with realistic ionic properties.” In: *Circulation research* 90.9 (2002).

[185] Robert Lemery et al. “Normal atrial activation and voltage during sinus rhythm in the human heart: An endocardial and epicardial mapping study in patients with a history of atrial fibrillation”. In: *Journal of Cardiovascular Electrophysiology* 18.4 (2007), pp. 402–408.

[186] Michael A. Colman et al. “In silico assessment of genetic variation in KCNA5 reveals multiple mechanisms of human atrial arrhythmogenesis”. In: *PLoS Computational Biology* 13.6 (June 2017). Ed. by Alexander V Panfilov, e1005587.

[187] Yingjing Feng et al. “Robust Atrial Ectopic Beat Classification From Surface ECG Using Second-Order Blind Source Separation”. In: *2020 Computing in Cardiology*. 2020, pp. 1–4.

[188] Cesare Corrado et al. “Using machine learning to identify local cellular properties that support re-entrant activation in patient-specific models of atrial fibrillation”. In: *Europace* 23.Supplement_1 (Mar. 2021), pp. I12–I20.

[189] Anand N. Iyer and Richard A. Gray. “Experimentalist’s approach to accurate localization of phase singularities during reentry”. In: *Annals of Biomedical Engineering* 29.1 (2001), pp. 47–59.

[190] Maurits A Allessie, F. I.M. Bonke, and F. J.G. Schopman. “Circus movement in rabbit atrial muscle as a mechanism of tachycardia. III. The ’leading circle’ concept: a new model of circus movement in cardiac tissue without the involvement of an anatomical obstacle”. In: *Circulation Research* 41.1 (1977), pp. 9–18.

[191] Nicholas Child et al. “Unraveling the Underlying Arrhythmia Mechanism in Persistent Atrial Fibrillation: Results from the STARLIGHT Study”. In: *Circulation: Arrhythmia and Electrophysiology* 11.6 (2018).

[192] Andrew Ng. “Machine learning yearning”. In: *URL: http://www. mlyearning. org/(96)* 139 (2017).

[193] Koichiro Kumagai et al. “A new approach for complete isolation of the posterior left atrium including pulmonary veins for atrial fibrillation”. In: *Journal of Cardiovascular Electrophysiology* 18.10 (2007), pp. 1047–1052.

[194] Jayakumar Sahadevan et al. “Epicardial mapping of chronic atrial fibrillation in patients: Preliminary observations”. In: *Circulation* 110.21 (2004), pp. 3293–3299.

[195] R. Sameni, C. Jutten, and M.B. Shamsollahi. “Multichannel Electrocardiogram Decomposition Using Periodic Component Analysis”. In: *IEEE Transactions on Biomedical Engineering* 55.8 (Aug. 2008), pp. 1935–1940.

[196] Emanuel Parzen. “On spectral analysis with missing observations and amplitude modulation”. In: *Sankhyā: The Indian Journal of Statistics, Series A* (1963), pp. 383–392.

[197] N. V. Chawla et al. “SMOTE: Synthetic Minority Over-sampling Technique”. In: *Journal of Artificial Intelligence Research* 16 (June 2002), pp. 321–357.

[198] Sudhir Varma and Richard Simon. “Bias in error estimation when using cross-validation for model selection”. In: *BMC Bioinformatics* 7 (2006).

[199] Guillaume Lemaître, Fernando Nogueira, and Christos K. Aridas. “Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning”. In: *Journal of Machine Learning Research* 18.17 (2017), pp. 1–5.

[200] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[201] Jiapu Pan and Willis J. Tompkins. “A Real-Time QRS Detection Algorithm”. In: *IEEE Transactions on Biomedical Engineering* BME-32.3 (1985), pp. 230–236.

[202] E. L. Kaplan and Paul Meier. “Nonparametric Estimation from Incomplete Observations”. In: *Journal of the American Statistical Association* 53.282 (June 1958), pp. 457–481.

[203] Cameron Davidson-Pilon. “lifelines: survival analysis in Python”. In: *Journal of Open Source Software* 4.40 (2019), p. 1317.

[204] R. S. MacLeod et al. “Effects of heart position on the body-surface electrocardiogram”. In: *Journal of Electrocardiology* 33 (2000), pp. 229–237.

[205] Giorgio Luongo et al. “Machine learning enables noninvasive prediction of atrial fibrillation driver location and acute pulmonary vein ablation success using the 12-lead ECG”. In: *Cardiovascular Digital Health Journal* 2.2 (2021), pp. 126–136.

[206] Angel Arenal et al. “Dominant frequency differences in atrial fibrillation patients with and without left ventricular systolic dysfunction”. In: *Europace* 11.4 (2009), pp. 450–457.

[207] Mohammad Hassan Shariat et al. “Regional Dominant Frequency: A New Tool for Wave Break Identification During Atrial Fibrillation”. In: *Frontiers in Cardiovascular Medicine* 5.June (2018), pp. 1–12.

[208] D Sculley. *Web-Scale K-Means Clustering*.

[209] J MacQueen. *Some methods for classification and analysis of multivariate observations*. Proc. 5th Berkeley Symp. Math. Stat. Probab., Univ. Calif. 1965/66, 1, 281-297 (1967). 1967.

[210] D. R. Cox. "Regression Models and Life-Tables". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 34.2 (Jan. 1972), pp. 187–202.

[211] Richard M Simon et al. "Using cross-validation to evaluate predictive accuracy of survival risk classifiers based on high-dimensional data". In: *Briefings in Bioinformatics* 12.3 (2011), pp. 203–214.

[212] Tim P. Morris et al. "Proposals on Kaplan-Meier plots in medical research and a survey of stakeholder views: KMunicate". In: *BMJ Open* 9.9 (2019).

[213] Peter J. Rousseeuw. "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis". In: *Journal of Computational and Applied Mathematics* 20 (Nov. 1987), pp. 53–65.

[214] N Mantel. "Evaluation of survival data and two new rank order statistics arising in its consideration". In: *Cancer chemotherapy reports* 50.3 (1966), 163—170.

[215] Masateru Takigawa et al. "Differentiating atrial tachycardias with centrifugal activation: Lessons from high-resolution mapping". In: *Heart Rhythm* 18.7 (2021), pp. 1122–1131.

[216] Hiroshi Morita et al. "The role of coronary sinus musculature in the induction of atrial fibrillation". In: *Heart Rhythm* 9.4 (2012), pp. 581–589.

[217] Enid Van Nieuwenhuyse et al. "Evaluation of Directed Graph-Mapping in Complex Atrial Tachycardias". In: *JACC: Clinical Electrophysiology* 7.7 (2021), pp. 936–949.

[218] Peter Langfield et al. "A novel method to correct repolarization time estimation from unipolar electrograms distorted by standard filtering". In: *Medical Image Analysis* 72 (Aug. 2021), p. 102075.

[219] Xingjian Shi et al. "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting". In: (2015), pp. 1–9.

[220] Tollef Struksnes Jahren et al. "Estimation of end-diastole in Cardiac Spectral Doppler Using Deep Learning". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 3010.c (2020), pp. 1–1.

[221] Zhaohan Xiong et al. "ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network". In: *Physiological Measurement* 39.9 (Sept. 2018), p. 094006.

[222] Younghyun Jo et al. "Deep Video Super-Resolution Network Using Dynamic Upsampling Filters Without Explicit Motion Compensation". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2018), pp. 3224–3232.

[223] Diederik P Kingma and Jimmy Ba. “Adam: A Method for Stochastic Optimization”. In: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* (Dec. 2014).

[224] Francois Rivest and Richard Kohar. “A New Timing Error Cost Function for Binary Time Series Prediction”. In: *IEEE Transactions on Neural Networks and Learning Systems* 31.1 (2020), pp. 174–185.

[225] Balvinder S. Handa et al. “Ventricular fibrillation mechanism and global fibrillatory organization are determined by gap junction coupling and fibrosis pattern”. In: *Cardiovascular Research* 117.4 (2021), pp. 1078–1090.

[226] Claudia Nagel et al. “A bi-atrial statistical shape model for large-scale in silico studies of human atria: Model development and application to ECG simulations”. In: *Medical Image Analysis* 74 (2021), pp. 1–14.

[227] Abien Fred Agarap. “Deep Learning using Rectified Linear Units (ReLU)”. In: (Mar. 2018).