
LEARNING TO PREDICT GLOBAL ATRIAL FIBRILLATION DYNAMICS FROM SPARSE MEASUREMENTS

A PREPRINT

Alexander Jenkins^{1,2†}, Andrea Cini^{3†}, Joseph Barker², Alexander Sharp⁴, Arunashis Sau²
Varun Valentine², Srushti Valasang², Xinyang Li², Tom Wong², Timothy Betts⁵
Danilo Mandic¹, Cesare Alippi^{3,6‡}, Fu Siong Ng^{2‡}

¹Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom

²National Heart and Lung Institute, Imperial College London, United Kingdom

³The Swiss AI Lab IDSIA USI-SUPSI, Università della Svizzera italiana, Switzerland

⁴Department of Engineering Science, University of Oxford, United Kingdom

⁵Oxford University Hospitals NHS Foundation Trust, United Kingdom

⁶Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Italy

Correspondence to: f.ng@imperial.ac.uk

ABSTRACT

Catheter ablation of Atrial Fibrillation (AF) consists of a one-size-fits-all treatment with limited success in persistent AF. This may be due to our inability to map the dynamics of AF with the limited resolution and coverage provided by sequential contact mapping catheters, preventing effective patient phenotyping for personalised, targeted ablation. Here we introduce FIBMAP, a graph recurrent neural network model that reconstructs global AF dynamics from sparse measurements. Trained and validated on 51 non-contact whole atria recordings, FIBMAP reconstructs whole atria dynamics from 10% surface coverage, achieving a 210% lower mean absolute error and an order of magnitude higher performance in tracking phase singularities compared to baseline methods. Clinical utility of FIBMAP is demonstrated on real-world contact mapping recordings, achieving reconstruction fidelity comparable to non-contact mapping. FIBMAP's state-spaces and patient-specific parameters offer insights for electrophenotyping AF. Integrating FIBMAP into clinical practice could enable personalised AF care and improve outcomes.

Keywords Atrial Fibrillation · electrophysiological mapping · catheter ablation · graph neural network · recurrent neural network · spatiotemporal · imputation

1 Introduction

Atrial Fibrillation (AF) is the most common cardiac arrhythmia with a lifetime risk of 1 in 3 [1]. AF is estimated to affect 37.5 million people worldwide - 0.5% of the global population - with a 60% projected rise by 2050 [2]. The arrhythmia underpins much of global morbidity and mortality as a major cause of stroke [3], heart failure [4] and death [5]. Global healthcare systems experience a significant and rising cost burden managing AF, with direct patient costs alone estimated to be 2.4% of the United Kingdom's £182 billion healthcare budget [6].

Catheter ablation is a common treatment for AF [7, 8] and aims to aid the maintenance of sinus (normal) rhythm through the controlled destruction of cardiac tissue via heating or freezing [9]. In this way, it is possible to electrically isolate regions of the atria that initiate or sustain AF. Pulmonary vein isolation, which isolates AF triggers in the pulmonary

[†]Joint first authors

[‡]Joint senior authors

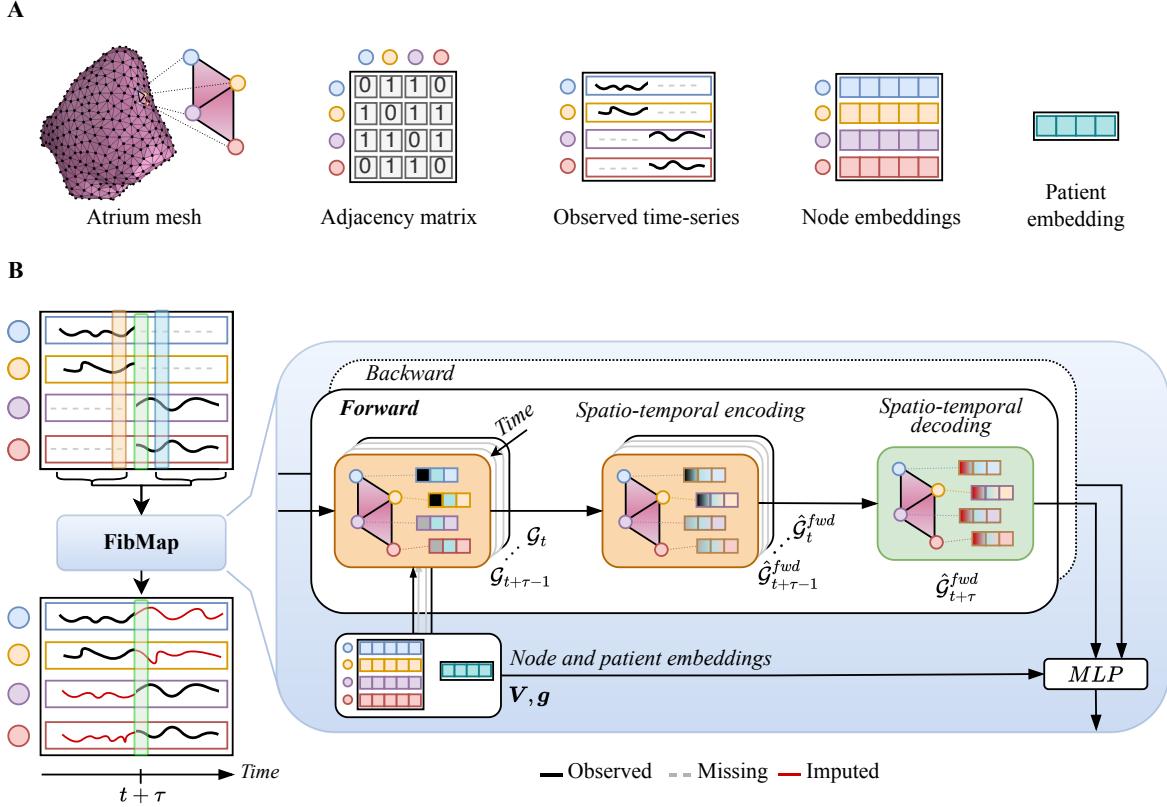


Figure 1: The inputs (A) and architecture (B) of FIBMAP. A) The atrium is discretised into nodes and edges via a triangulated mesh, and from this the graph adjacency matrix (describing the coupling between nodes) and the observed time series are derived. The patient-specific parameters (node and patient embeddings) are also provided as input to personalise FIBMAP’s imputation maps. B) FIBMAP is instantiated as a bidirectional graph recurrent neural network (GRNN), a state-space model that learns a representation for each node by propagating information forwards and backwards through space and time. Node and patient embeddings are provided to a multilayer perceptron (MLP) decoder, which personalises how the learnt representation is mapped to a signal value.

veins from the left atrium, is central to AF ablation, yet its success rates are as low as 50% at 5 years in persistent AF, with effectiveness diminishing over time [10]. Other ablation targets, such as linear lesions to compartmentalise the atria and targeted ablation of putative AF drivers have not significantly improved outcomes [11, 12, 13, 14, 15]. These approaches fail to tailor interventions to the unique electrophenotypes within the spectrum of AF observed in patients, leaving a one-size-fits-all treatment plan for a highly heterogeneous disorder [16].

The major limitation in applying tailored ablation strategies within the invasive electrophysiology laboratory is our inability to map AF effectively, that is, reconstruct AF dynamics on the surface of the heart. Sequential contact mapping is routinely used to record electrical signals from the atria, wherein a multipolar electrode catheter is inserted into the atrial chamber and placed in contact with the endocardial surface during the invasive procedure. The chamber is scanned sequentially with voltage recorded as a function of time and position, and the results are computationally stitched together. This approach is powerful for organised arrhythmias, such as focal or re-entrant atrial tachycardias, where the regularity of the dynamics allows for accurate stitching and thereby precise localisation of the arrhythmic source [17], resulting in a curative targeted ablation procedure. However, this approach is ineffective in AF due to its highly disorganized nature, with beat-to-beat variations in wavefront propagation, meaning the non-continuous recordings cannot be sensibly stitched together. While there have been attempts at continuous whole atria mapping, these typically require different modalities such as non-contact catheters that are not routinely used in the clinic since they are expensive, often lack the spatial resolution and coverage necessary to effectively target AF drivers, and increase procedural complexity and risk [14, 18, 19, 20].

The novel approach outlined here is to consider the atria in AF as a network of oscillators coupled through a graph structure, where non-overlapping regions of tissue are represented as nodes and their coupling to neighbouring tissue

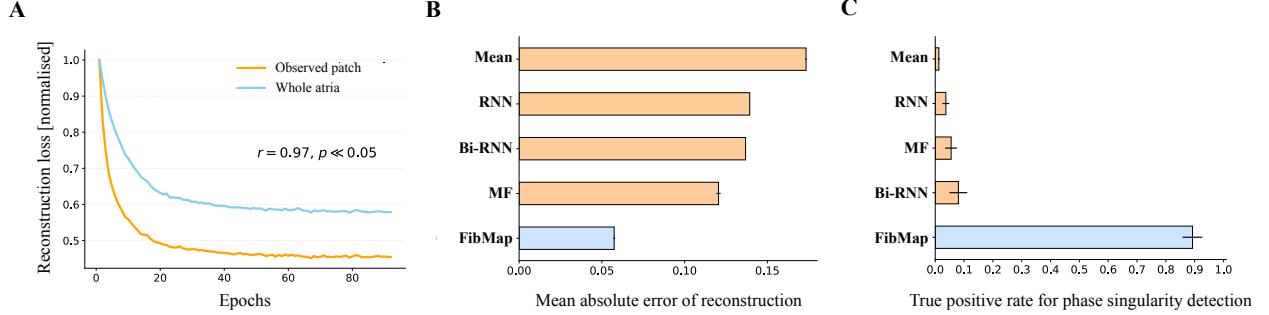


Figure 2: Quantitative results of FIBMAP imputation mapping. A) Reconstruction loss as a function of the number of epochs for the fine-tuning of FIBMAP on the validation set, with a strong correlation of $r = 0.97$ ($p < 0.0001$) present between the loss curves of the observed patch and whole atria. B) Test set reconstruction performance of all models quantified using the mean absolute error (MAE) across all space and time. C) Test set performance for detecting phase singularities (PSs), quantified using the true positive rate of their detection.

via edges [21, 22, 23]. In doing so, graph neural networks (GNNs) [24, 25, 26], the generalisation of deep learning methods to irregular (non-Euclidean) data, can be applied to model the dynamics of AF and reconstruct them from measurements of sequential contact mapping. This method, which we term *imputation mapping*, aims to reconstruct the dynamics of AF via imputation. By leveraging a graph recurrent neural network (GRNN) [27, 28], a non-linear state-space model suited for systems of coupled oscillators and spatiotemporal data analysis, our model, FIBMAP, offers a solution for imputation mapping that can 1) express the rich dynamics of fibrillation for each patient, 2) effectively reconstruct the AF dynamics from the sparse and routinely collected sequential contact measurements, and 3) efficiently generalise to new patients for use in a clinical setting. To this end, imputation mapping is designed to reconstruct whole atria dynamics from routine sequentially collected contact mapping measurements, resulting in an integrative solution for AF mapping that holds the potential to unlock personalised AF care, enhance clinical outcomes, and reduce the duration, complexity, and risk of the AF mapping procedure.

2 Results

2.1 FIBMAP performs accurate whole atria imputation mapping for AF

Leveraging spatiotemporal message passing [28] on a dynamical graph representation of the input, FIBMAP effectively propagates information from valid measurements across time and the atrial surface to accurately perform imputation mapping. Our solution personalises the reconstruction of AF dynamics through unique node and patient embedding parameters, akin to word embeddings in natural language processing [29]; conversely, shared parameters are utilised across patients to capture common elements of the dynamics such as the physics of wave propagation. The input representation and architecture of FIBMAP are detailed in Figure 1.

Our dataset comprises non-contact recordings from 51 patients with persistent AF, obtained during an AF ablation procedure using the AcQMap system before pulmonary vein isolation. The AcQMap system uses non-contact electrodes to sense intra-cavitary unipolar electrograms, from which dipole density measurements are inversely derived to provide improved spatial resolution of electrical activity on the atrial surface [30, 31]. Each recording simultaneously samples approximately 3500 nodes across the entire atria at 3000 Hz for 5-20 seconds. Patients were treated between 2016 and 2023 at two centres in the United Kingdom: The Royal Brompton Hospital, London, and the John Radcliffe Hospital, Oxford. The cohort (mean age 64 ± 11 years, 69% male) had a mean BMI of $29 \pm 5 \text{ kg/m}^2$, with comorbidities including hypertension (39%), heart failure (33%), and diabetes mellitus (10%) (see Table S.T1 in supplementary materials for full details). The procedure involved left atrial mapping, followed by pulmonary vein isolation, and in some cases further adjunctive linear ablations (anterior, posterior, septal) and cardioversion to restore sinus rhythm. These recordings provide a ground truth of human AF dynamics with fidelity suitable for developing and validating FIBMAP. After resampling these signals spatially to 500 nodes, temporally to 70 Hz, and normalising the signals to enable consistent analysis of the underlying spatiotemporal patterns independent of measurement units, the dataset was stratified to ensure a spectrum of AF dynamics within training (70%), validation (10%), and test (20%) sets.

Continuous whole atria mapping with systems such as AcQMap are not routinely used in the clinic. Instead, contact mapping catheters are used to precisely measure electrical signals from the atria, with the limitation that they provide sparse measurements across the surface. To emulate routine clinical data collection within the invasive electrophysiology

Table 1: Imputation results for 10% observed area, dwell time of 1 second, and no spatial overlap.

Model	Metrics				
	MAE	MSE	MRE	MAPE	PS TPR
Mean	0.1734 ± 0.0005	0.0507 ± 0.0003	35.0833 ± 0.1043	47.7258 ± 0.3260	0.0126 ± 0.0035
Univariate RNN	0.1393 ± 0.0001	0.0298 ± 0.0000	28.1724 ± 0.0213	44.0633 ± 0.2098	0.0369 ± 0.0121
Univariate Bi-RNN	0.1368 ± 0.0001	0.0290 ± 0.0001	27.6553 ± 0.0251	42.8590 ± 0.1856	0.0803 ± 0.0309
MF	0.1205 ± 0.0012	0.0254 ± 0.0005	24.3715 ± 0.2434	32.5598 ± 0.5609	0.0552 ± 0.0207
FIBMAP	0.0574 ± 0.0005	0.0069 ± 0.0001	11.5868 ± 0.1098	16.7715 ± 0.4589	0.8924 ± 0.0342

laboratory, a sequential contact mapping strategy was simulated from the whole atria non-contact recordings as a self-avoiding walk of the multipolar catheter (represented as a patch of observations) on the atrial surface; whereby the catheter surface area, dwell time (duration of recording at each location), and spatial overlap, could be controlled (see Figure 3A). During training, FIBMAP was optimised for imputation mapping through a whole atria reconstruction loss and a self-supervised learning approach, wherein various multipolar catheter paths were sampled, and parameters augmented (see subsection S.3.3 of supplementary materials for more details). As a result, the pre-trained FIBMAP model, adept at reconstructing a spectrum of AF dynamics from sparse measurements, remains robust against variations in multipolar mapping.

This model is efficiently adapted to new patients through a fine-tuning transfer learning procedure, designed to enable accurate whole atria reconstruction from sequential contact mapping data. Our fine-tuning approach preserves essential knowledge for whole atria reconstruction acquired during training by fixing model parameters, while quickly personalising the model by optimising only the patient-specific parameters using an observed patch reconstruction loss (see subsection S.3.3 of supplementary materials for more details). The fine-tuning procedure was configured on the validation set, whereby a strong correlation (Pearson’s $r = 0.97$, $p < 0.0001$) between the observed patch and whole atria reconstruction loss curves was found (see Figure 2A). This indicates that our fine-tuning procedure enables accurate whole atria reconstruction and generalises effectively to new patients without overfitting to the signals in the observed patches. The performance of FIBMAP imputation mapping on new and unseen patients, through our fine-tuning procedure, was quantified on the patients in the test set. Testing involved simulating sequential contact mapping for test set patients with a catheter surface area of 10% of the atrium, a dwell time of 1 second, and no spatial overlap between successive patches (as shown in Figure 3A). FIBMAP was fine-tuned using the configuration found in validation, with the observed patch reconstruction loss monitored to perform early stopping. Test set fine-tuning took a total of 3 hours and 40 minutes, averaging just 22 minutes per patient. After fine-tuning, FIBMAP’s performance was confirmed through various metrics against ground truth whole atrium signals, all computed on the normalised signals. FIBMAP significantly outperformed baseline models in predicting whole atrium signals from simulated sequential contact mapping measurements. Figure 2B illustrates these performances, with the complete set of results shown in Table 1. Specifically, FIBMAP achieved an mean absolute error (MAE) of 0.0574 ± 0.0005 , a 2.1x improvement compared to the best baseline imputation model, matrix factorisation (MF), which had an MAE of 0.1205 ± 0.0012 . In mean squared error (MSE) and mean absolute percentage error (MAPE), FIBMAP scored 0.0069 ± 0.0001 and 16.7715 ± 0.4589 , respectively, compared to MF’s 0.0254 ± 0.0005 and 32.5598 ± 0.5609 .

In addition, the performance of the imputation models was assessed by their ability to track the phase singularities (PSs) present in the recordings, a characteristic of fibrillation dynamics [32]. These points, where the phase of electrical activity is undefined, mark the tips of rotating spiral waves that could drive and maintain AF, representing potential ablation targets and providing crucial insights into the underlying fibrillation mechanisms. Two independent observers manually labelled the AFs in the central 1-second segment (70 frames) of each map from the different imputation models and patients. FIBMAP excelled in tracking PSs, achieving a true positive rate (TPR) of 0.8924 ± 0.0342 , an $11.1 \times$ improvement upon the next closest competitor TPR of 0.0803 ± 0.0309 , as illustrated in Figure 2C and Table 1. These results demonstrate FIBMAP’s capability to accurately reconstruct AF dynamics across space, time, and patients, significantly improving upon existing imputation methods.

Our empirical results confirm FIBMAP’s capability to reconstruct the complex dynamics of AF accurately. Figure 3B illustrates FIBMAP’s signal reconstructions at individual nodes compared to baseline methods (columns) across two test patients (rows), highlighting its ability to reconstruct the temporal dynamics of AF. While small amplitude variations are observed between FIBMAP’s reconstruction and the ground truth recordings, the true signal often resides within FIBMAP’s predicted confidence intervals (90th – 10th quantiles, see subsection S.2.3 of supplementary materials for more details). Furthermore, FIBMAP’s reconstruction maintains phase coherence with the ground truth recordings, allowing for the accurate visualisation of phase maps across the atrium.

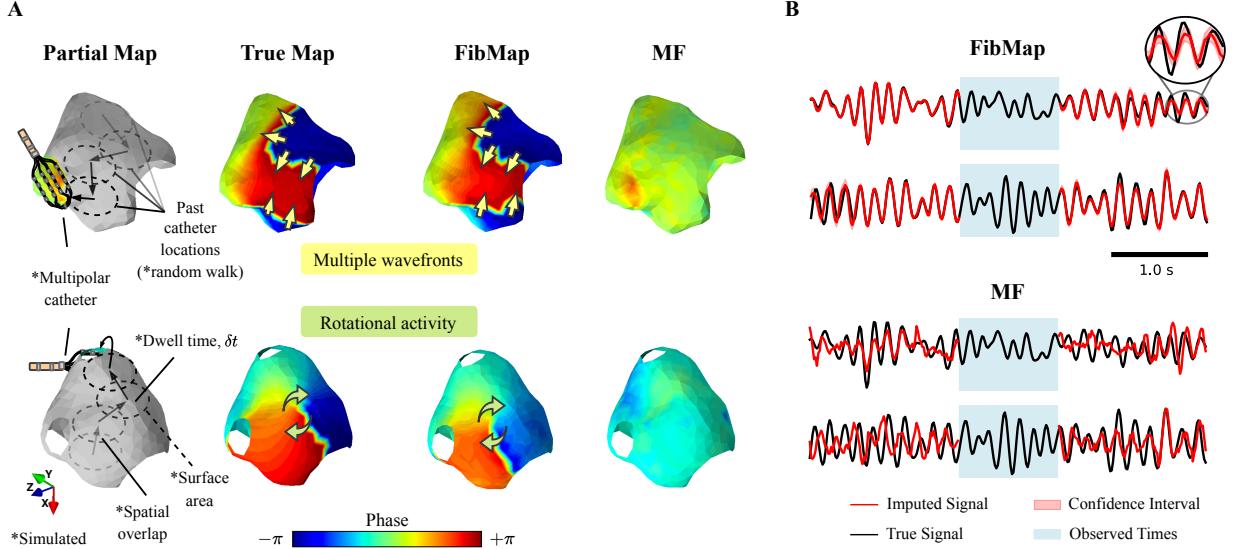


Figure 3: Qualitative results of FIBMAP imputation mapping. A) Snapshots of the imputation maps of FIBMAP and MF versus the ground truth AcQMap phase maps for two different patients (rows). FIBMAP imputes the missing signals (grey regions) from sparse observations (10% of atria) of AcQMap recordings with a simulated sequential contact mapping multipolar catheter (surface area of 10%, no spatial overlap and 1 second dwell time). B) Node-level temporal reconstructions for two different patients (rows) for FIBMAP and MF. Both the maps and signal traces demonstrate the superior performance of FIBMAP.

Figure 3A provides a snapshot of FIBMAP’s phase maps derived from the predicted signal, demonstrating its precise reconstruction of AF spatial dynamics for two different patients (rows). These maps reveal characteristic features of complex AF dynamics such as multiple wavefronts and rotational activity that were otherwise not visible in the observed patches. In comparison, our empirical findings in Figure 3B underscore the limitations of the baseline models. While MF could in principle be used here, it fails in capturing the complex spatiotemporal dynamics that characterise AF, as evidenced by reconstruction errors in both amplitude and phase in Figure 3B which result in poor phase map reconstruction in Figure 3A.

2.2 FIBMAP reconstructs personalised AF dynamics from clinical sequential contact mapping data

FIBMAP’s ability to reconstruct whole atria dynamics from real-world sequential contact mapping data was validated retrospectively using recordings from the Advisor™ HD Grid Mapping catheter (16 electrodes arranged in a grid and evenly spaced 3mm apart, referred to as HD Grid) and the EnSite Precision mapping system (Abbott Laboratories, Lake County, Illinois, United States). This system is routinely used for mapping cardiac arrhythmia and guiding ablation procedures in the invasive electrophysiology laboratory. For three AF patients in the test set, both EnSite Precision contact mapping and non-contact AcQMap recordings of different durations (tens of minutes vs. seconds) were collected in sequence (non-contemporaneously) before ablation was performed. EnSite Precision HD Grid recordings were resampled to 500 nodes at 70 Hz, electrograms were normalised, and imputation maps were created from the EnSite Precision HD Grid using our fine-tuning procedure (see subsection S.3.7 of supplementary materials for more details).

A comparative approach was developed to assess FIBMAP imputation maps from the EnSite Precision HD Grid against the whole atria ‘ground truth’ provided by AcQMap. The non-contemporaneous nature of these recordings and their varying durations present analytical challenges. AF wavefront patterns vary beat-to-beat, preventing direct temporal alignment and direct comparison between the imputed maps from the EnSite Precision HD Grid versus the ground truth of the global AcQMap recordings (which were collected minutes apart). To overcome this, a sliding window analysis was implemented (Figure 4A) to cross-correlate segments of ground truth AcQMap and imputed FIBMAP phase signals, quantifying the overall dynamical similarity between these non-contemporaneous recordings, without requiring exact temporal matching. This approach enables fair and meaningful comparison between the two data modalities (imputed maps vs. ground truth). The resulting cross-correlation distributions were analysed to determine whether FIBMAP captures patient-specific dynamics and distinguishes genuine electrophysiological patterns from arbitrary correlations expected by random chance.

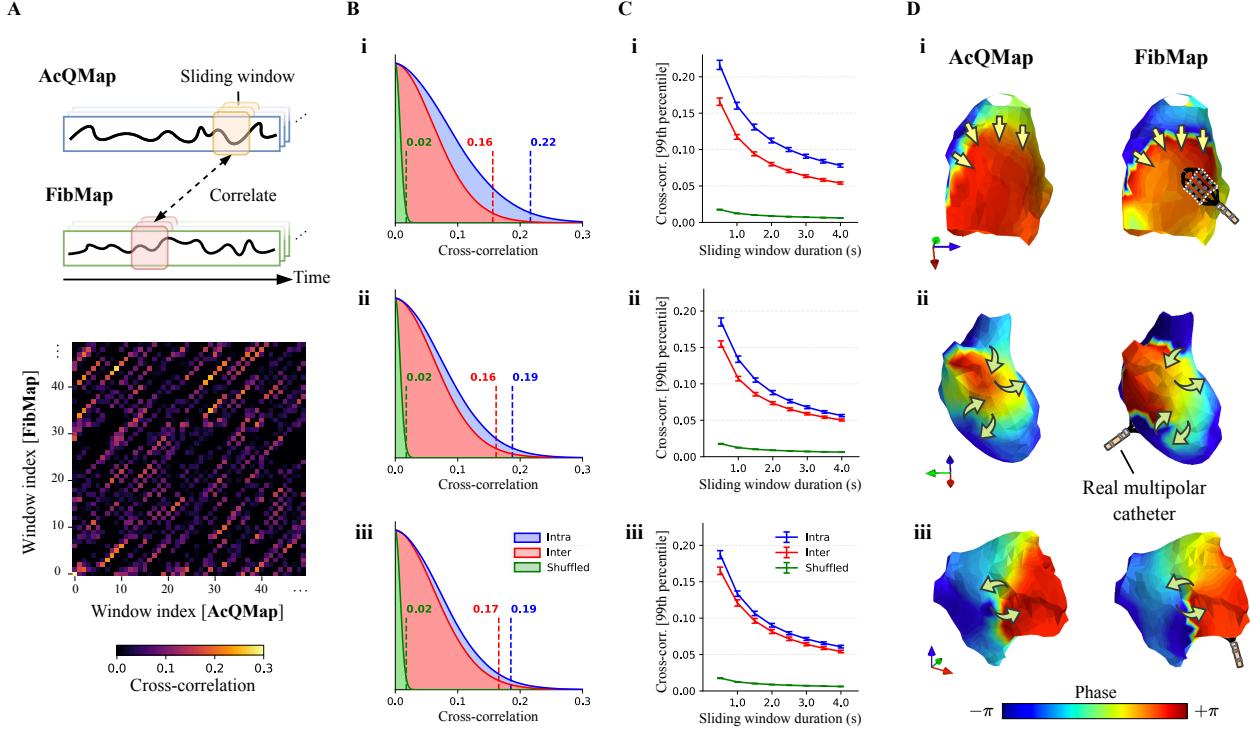


Figure 4: Validation of FIBMAP imputation maps from EnSite Precision HD Grid Mapping against non-contemporaneous ground truth AcQMap recordings. A) Sliding window cross-correlation analysis between AcQMap and FIBMAP phase signals, enabling comparison between non-contemporaneous recordings by measuring the similarity between AF patterns across different time windows. The correlation matrix below shows pairwise comparisons between all possible window combinations. For three patients (indexed i-iii), B) shows the kernel density estimates of pairwise cross-correlations computed between AcQMap and FIBMAP maps from the same patient (intra), different patients (inter), and spatiotemporally shuffled maps (random baseline), using 0.5-second sliding windows. Vertical dashed lines indicate the 99th percentile, with consistently higher values for intra-patient (0.19-0.22) versus inter-patient (0.16-0.17) correlations and shuffled baseline (0.02), demonstrating patient-specific pattern capture. C) Temporal robustness analysis showing the 99th percentile of cross-correlations against sliding window duration. Non-overlapping confidence intervals (computed via bootstrap sampling) between intra-patient and other distributions confirm that FIBMAP can capture patient-specific dynamics across different temporal scales. D) Representative snapshots of phase maps from highly correlated windows, comparing AcQMap recordings with FIBMAP maps derived from real HD Grid catheter measurements covering < 10% of the atrial surface per time. The maps demonstrate the ability of FIBMAP to capture AF dynamics which are not directly visible in the original HD Grid measurements, including multiple wavefronts and rotational patterns.

To investigate patient specificity, distributions of cross-correlation were computed by comparing AcQMap recordings (ground truth) to FIBMAP (imputed maps) of the same (intra-) patient and different (inter-) patients using a sliding window of duration 0.5 seconds. To compute the cross-correlation, recordings were projected between source and target surfaces using a nearest-neighbours approach (see subsection S.3.4 of supplementary materials for more details). Cross-correlation distributions were plotted as kernel density estimates in Figure 4Bi-iii for all three patients, revealing consistently wider tails for intra-patient correlations between AcQMap and FIBMAP, with 99th percentiles ranging from 0.19-0.22 compared to 0.16-0.17 for inter-patient correlations. The higher frequency of strong correlations within the same patient demonstrates FIBMAP’s ability to capture patient-specific dynamics rather than generic patterns common across patients.

To assess whether these correlations were meaningful, we compared them against a random baseline created by spatiotemporally shuffling the FIBMAP imputation maps and then performing an intra-patient comparison with AcQMap. The significant separation between intra-patient and random distributions (99th percentiles shown by vertical dashed lines in Figure 4Bi-iii) confirms that the similarities between AcQMap and FIBMAP are not due to chance, validating FIBMAP’s ability to impute AF dynamics. If this were due to chance, the observed 99th percentile for

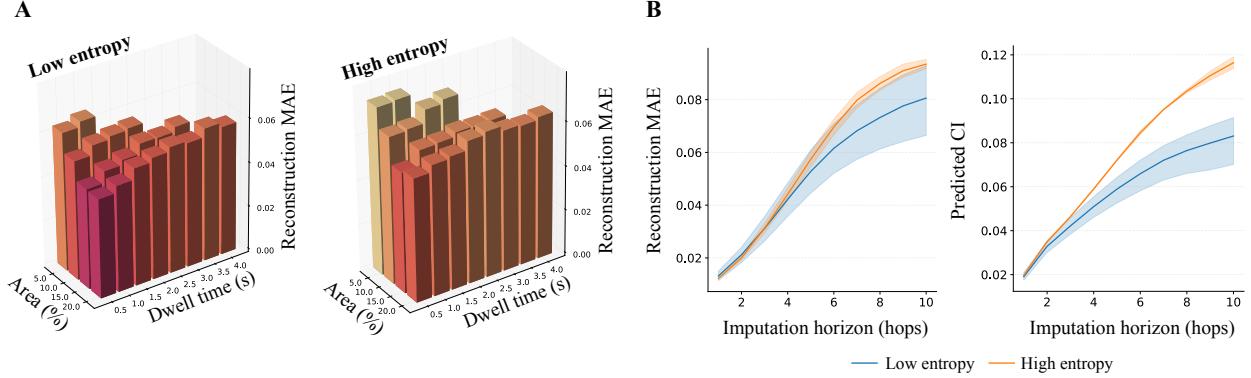


Figure 5: Results of FIBMAP sensitivity analysis. A) MAE of FIBMAP reconstructions as a function of catheter surface area, dwell time and entropy (left vs. right). B) Reconstruction MAE of FIBMAP and the predicted confidence intervals as a function of imputation horizon.

intra-patient correlations would lie at approximately 0.02 for each patient, an order of magnitude lower than the observed values of 0.19-0.22 (see Figure 4Bi-iii).

To test the robustness of these findings, sliding window durations were varied from 0.5 to 4.0 seconds (Figure 4Ci-iii). The 99th percentiles of the cross-correlation distributions were plotted against window duration, showing that while these values generally decreased with longer windows, the 99th percentiles for intra-patient correlations consistently remained higher than both inter-patient and shuffled comparisons. For each patient, confidence intervals were computed via bootstrap sampling (see subsection S.3.7 of supplementary materials for more details), revealing non-overlapping intervals between intra-patient, inter-patient, and shuffled distributions across all window durations, confirming the statistical significance of these differences. Representative phase maps (Figure 4D) visually confirm the agreement between FIBMAP imputations and AcQMap ground truth recordings to reveal complex AF dynamics such as multiple wavefronts and rotational activity, which were otherwise not visible by the EnSite Precision HD Grid which covers less than 10% of the atrial surface.

These findings demonstrate that FIBMAP, through our fine-tuning procedure, can reconstruct personalised whole atria AF dynamics from real examples of multipolar catheter contact recordings collected in the invasive electrophysiology laboratory with fidelity comparable to AcQMap. This demonstrates the capacity of imputation mapping to integrate with and significantly enhance the capabilities of routinely used multipolar catheter mapping systems.

2.3 The sensitivity of FIBMAP to variations in multipolar mapping and AF organisation

Returning to our analysis of simulated sequential contact mapping, a sensitivity analysis was conducted to assess FIBMAP’s practical application by evaluating its reconstruction performance of whole atria dynamics across various catheter areas, dwell times, and levels of AF organisation. Spatial overlap was fixed at zero for this analysis. AF organisation was quantified by computing the average Shannon entropy across the ground truth signals of each patient, where lower values indicate more organised dynamics. Patients were then divided into two groups of low and high entropy based on the median average Shannon entropy value.

Figure 5A illustrates the MAE of the whole atria reconstruction for both groups across varying catheter areas and dwell times. The low entropy group (more organised forms of AF) consistently exhibited lower MAE values, with 0.065 on average compared to 0.071 of the high entropy group. Furthermore, both groups demonstrated improved reconstruction performance with increased catheter surface area and reduced dwell time, achieving a minimal MAE of 0.045 in the low entropy group with a 20% catheter surface area and a 0.5s dwell time. A shorter dwell time enables faster sampling of the whole atrium, allowing it to be covered multiple times within the finite duration of our recordings. For each combination of dwell time and catheter surface area depicted in Figure 5A, we ensure that each patient’s atrium is covered at least once. Combinations that do not meet this criterion are not shown. While these results could guide the optimal use of FIBMAP, we were unable to evaluate its performance for longer durations due to the limited length of our recordings. However, in practice, sampling duration is not restricted.

Performance was assessed based on the imputation horizon to reveal insights into FIBMAP’s capabilities and their implications for practical application. The imputation horizon, defined as the spatiotemporal distance from the closest observed node and time point (measured in hops on a spatiotemporal graph), measures how far (in space and time)

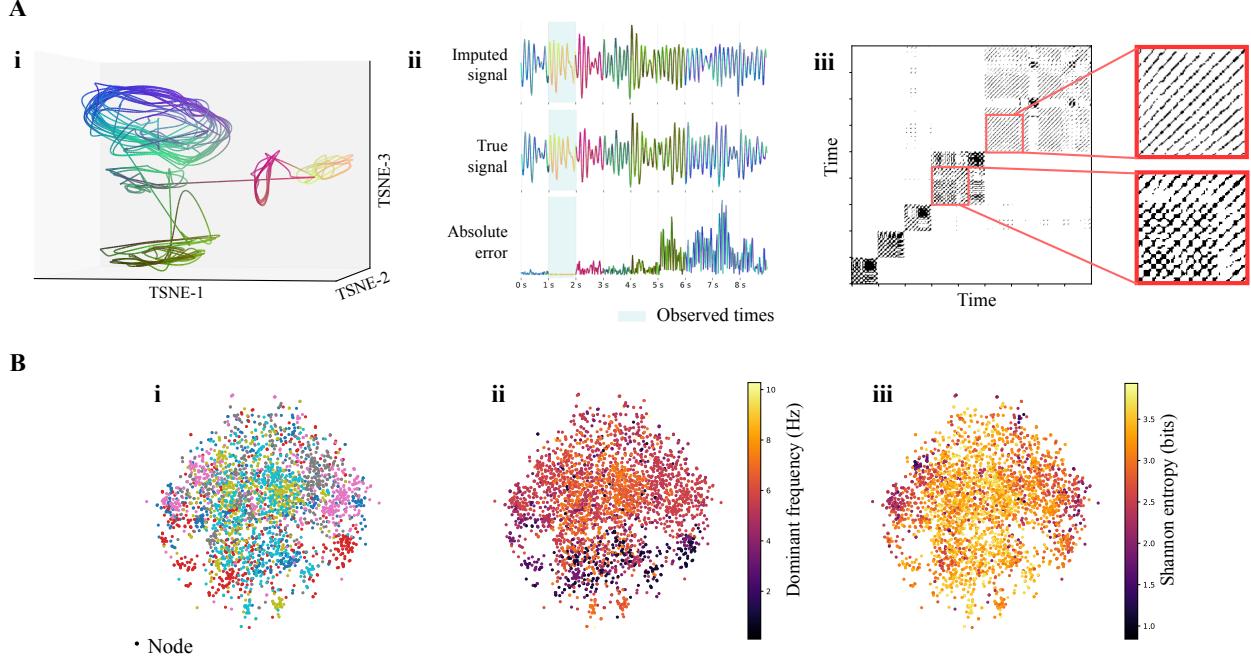


Figure 6: Interpretation of FIBMAP hidden state and patient-specific node parameters. Ai) Trajectory of the hidden state of a node in the test set, where the state space has been reduced to 3D using t-SNE dimensionality reduction. Trajectories are coloured by their position in the state space by converting the 3D coordinate to RGB. Aii) The imputed and true signals are plotted for the same node, coloured by their position in the state space. The absolute error between the imputed and true signals is also shown. Aiii) A recurrence plot summarising the state-space trajectory, where different dynamical structures are observed over time. B) Node embeddings in the test set coloured by (i) patient, (ii) dominant frequency and (iii) Shannon entropy.

FIBMAP can reliably predict beyond the observed data points. Figure 5B demonstrates that the high entropy group experienced a sharper increase in MAE with increasing imputation horizon compared to the low entropy group, where MAE eventually plateaued, possibly due to FIBMAP’s predictions losing phase coherence with the ground truth. Concurrently, Figure 5C demonstrates a corresponding increase in predicted confidence intervals with the imputation horizon. This underscores FIBMAP’s ability to quantify uncertainty in long-range predictions, which is crucial for reliably interpreting whole atria imputation maps and making informed clinical assessments.

2.4 Revealing structure in FIBMAP’s state- and embedding-spaces

FIBMAP’s design as a non-linear state-space model enables the visualisation of hidden state dynamics across all nodes and times. These hidden states exhibit chaotic dynamics, characterised by diverging trajectories around fixed points in state space, effectively capturing the inherent chaos of AF. In Figure 6Ai, we use t-SNE dimensionality reduction [33] to visualise the state-space trajectory of a given node in a 3D space. Converting these 3D coordinates to RGB values creates a colour-coding that reveals orbital patterns and transitions across different regions of state space.

To examine the relationship between state space regions and signal dynamics, Figure 6Aii shows the imputed and ground truth signals coloured by their state space positions, along with their absolute error to reveal state-dependent variations. While both state duration and absolute error vary, the dynamics show subtle changes between states. Figure 6Aiii presents a recurrence plot [34] of the state-space trajectory, revealing structured dynamics within and between states. For example, parallel lines to the main diagonal appear at later times, which is a characteristic of deterministic dynamics where trajectories repeatedly progress through the same sequence of states.

While a comprehensive analysis of state space dynamics exceeds the scope of this work, our findings demonstrate that FIBMAP’s formulation as a state-space model inherently captures important information about AF dynamics. Moreover, non-linear dynamics analysis tools, such as recurrence plots, can extract additional insights from FIBMAP. These insights could enable electrophenotyping of AF patients into broad subgroups, help predict the probability of benefit from ablation, and guide ablation strategy.

The learned patient and node embedding parameters in FIBMAP provide additional structural insights. Figure 6B visualises the node embeddings for each patient in the test set in 2D space using t-SNE dimensionality reduction [33]. Figure 6Bi shows each node (marker) coloured uniquely for each patient, revealing both localised clusters and nodes distributed across the space, suggesting similarities (and differences) between the nodes of different patients. To interpret these embeddings, Figure 6Bii and Figure 6Biii show the dominant frequency and Shannon entropy of the signal at each node, respectively. The node embeddings capture this information in a complex manner, with dominant frequency generally decreasing toward the bottom and Shannon entropy increasing toward the centre. These node embedding spaces could enable a systematic comparison of tissue properties across patients and anatomical locations, while patient embeddings could identify individuals with similar global dynamics.

These findings indicate that both FIBMAP’s hidden states and embedding parameters could aid in electrophenotyping patients. However, since follow-up information or spatial properties of the tissue were not available for our dataset, further research is needed to fully interpret and validate the information contained within the patient and node embeddings.

3 Discussion

The inability to map AF effectively in the invasive electrophysiology laboratory limits the application of tailored ablation strategies and potentially hinders curative treatment outcomes. To address this problem, we introduce imputation mapping to reconstruct full AF dynamics from recordings collected through non-continuous sequential contact mapping, which observe only a fraction of the atrium at a time. Until now, these sparse measurements could not be effectively stitched into a global map due to the highly disorganised nature of AF. To this end, our solution, termed FIBMAP, employs a novel GRNN-based non-linear state-space model for personalised reconstruction of global AF dynamics.

FIBMAP was trained and rigorously validated on a dataset of 51 persistent AF patients, using continuous non-contact whole atria recordings from the AcQMap system as ground truth and simulated sequential contact mapping to replicate clinical data collection. Our results show that FIBMAP outperforms the baseline imputation models in whole atria reconstruction, achieving over $2\times$ improvement in MAE and more than an order of magnitude improvement in PS detection rates, while observing only 10% of the atrium. Empirical results demonstrate that FIBMAP accurately captures both amplitude and phase relations across the atria, reconstructing complex features of AF dynamics such as rotational activity and multiple wavefronts. Beyond simulated contact mapping, we showed that FIBMAP successfully transfers to real clinical examples of sequential contact mapping, collected with the multipolar HD Grid catheter, to reconstruct AF dynamics with fidelity comparable to global AcQMap. This demonstrated the ability of FIBMAP to integrate sparse contact recordings of AF into a coherent imputation map. We expect further improvements in reconstruction quality when the catheter placement is guided prospectively (possibly by the results of our sensitivity analysis). Finally, visualisation of FIBMAP’s hidden state and embedding spaces reveals significant structure, suggesting potential for future electrophenotyping studies.

In summary, FIBMAP provides a comprehensive solution for imputation mapping by: 1) capturing the rich dynamics of fibrillation for each patient, 2) effectively reconstructing whole atria maps from sequential contact measurements, 3) producing uncertainty estimates to facilitate a clinical interpretation, 4) efficiently generalising to new patients for clinical use through our fine-tuning procedure, and 5) parameterising AF phenotypes, offering the opportunity for downstream work to investigate its potential in personalising AF care.

While methods for whole atria mapping do exist, such as the AcQMap system and non-invasive electrocardiographic imaging, these approaches have not demonstrated sufficient efficacy for routine clinical use, due in part to their non-contact nature [14, 18, 19, 20]. These non-contact methods suffer from having to calculate the electrograms on the heart surface rather than recording true contact electrograms. As a result, global non-contact mapping approaches suffer from low spatial resolution, leading to incorrect interpretation of AF dynamics [20]. Imputation mapping offers a novel approach to obtaining whole atria maps from routinely collected sequential contact mapping data, efficiently integrating with existing mapping multipolar catheters (such as EnSite Precision HD Grid) to reconstruct AF dynamics in a real-time or post hoc manner. Although our work has demonstrated a proof-of-concept for imputation mapping, future clinical studies should focus on validating its efficacy in guiding personalised ablative strategies. These studies should also determine the optimal spatial and temporal resolution of the imputation map for accurate identification of fibrillation drivers and their precise ablation.

A crucial design choice in our study was the use of data from the AcQMap system. Non-contact data provide a realistic ground truth for human AF, which is essential for accurate model training and validation. Using these measurements, FIBMAP can learn an empirical model of AF dynamics, ensuring an accurate representation of human AF. In contrast, using simulated whole atria signals would bias the model towards simplified and debated mathematical models of fibrillation [35, 36]. While this approach proved effective for our proof-of-concept validation on AcQMap and EnSite

Precision HD Grid recordings, if future work did identify limitations, FIBMAP can be trained on a mixture of real and simulated fibrillation data to enhance its robustness. However, we emphasise that the current design of FIBMAP incorporates a pre-trained model component that permits continual learning, improving accuracy over time as it is applied to and fine-tuned on contact mapping recordings from new patients across the spectrum of AF. Imputation mapping, as demonstrated by FIBMAP, offers a promising approach for reconstructing whole atria dynamics from routinely collected sparse sequential contact mapping data.

Future research will focus on proving the efficacy of imputation mapping in guiding personalised catheter ablation procedures and determining the optimal resolution requirements for its successful clinical deployment. Additionally, refinements to the architecture of FIBMAP will further improve imputation accuracy. For example, hierarchical architectures [37, 38] would improve the imputation performance at longer horizons. By effectively bridging the gap between sequential contact mapping and continuous whole atria mapping systems, imputation mapping allows clinicians to *see more with less*. Its integration with existing contact mapping systems has the potential to transform our interpretation of complex AF dynamics, ultimately improving the precision and effectiveness of AF ablation.

Acknowledgments

AJ was supported by the UKRI CDT in AI for Healthcare <http://ai4health.io> (Grant No. P/S023283/1). JB was supported by the British Heart Foundation (FS/CRTF/24/24624). AS received a research grant and travel support from Acutus Medical. TB received honoraria from Acutus Medical. CA was partly funded by the Swiss National Science Foundation under grant 204061: High-Order Relations and Dynamics in Graph Neural Networks. FSN was supported by the British Heart Foundation (RG/F/22/110078, RE/19/4/130023) and the National Institute for Health Research Imperial Biomedical Research Centre.

Author contributions

AJ conceptualised the solution, designed and performed the experiments, and wrote the manuscript. AC conceptualised the solution, designed the experiments, and edited the manuscript. JB designed the experiments and wrote and edited the manuscript. VV and SV manually annotated the phase singularities present in the recordings. A Sharp, A Sau, XL, TW, and TB supplied clinical data for analysis. DM, CA and FSN co-supervised the design of the solution and the empirical evaluation and edited the manuscript.

Competing interest declaration

AJ, AC, DM, CA and FSN are applicants on the patent: *A Method of Constructing Maps of Dynamical Variables on a Cardiac Surface* (UK Patent Application No. 2500830.1). All other authors declare no conflicts of interest.

References

- [1] Liping Mou et al. “Lifetime risk of atrial fibrillation by race and socioeconomic status: ARIC study (Atherosclerosis Risk in Communities)”. *Circulation: Arrhythmia and Electrophysiology* 11.7 (2018), e006350.
- [2] Giuseppe Lippi, Fabian Sanchis-Gomar, and Gianfranco Cervellin. “Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge”. *International Journal of Stroke* 16.2 (2021), pp. 217–221.
- [3] Aristeidis H. Katsanos, Hooman Kamel, Jeff S. Healey, and Robert G. Hart. “Stroke prevention in atrial fibrillation: Looking forward”. *Circulation* 142.24 (2020), pp. 2371–2388.
- [4] William H. Maisel and Lynne Warner Stevenson. “Atrial fibrillation in heart failure: Epidemiology, pathophysiology, and rationale for therapy”. *The American Journal of Cardiology* 91.6 (2003), pp. 2–8.
- [5] Philip A. Wolf, Janet B. Mitchell, Colin S. Baker, William B. Kannel, and Ralph B. D’Agostino. “Impact of atrial fibrillation on mortality, stroke, and medical costs”. *Archives of Internal Medicine* 158.3 (1998), pp. 229–234.
- [6] S Stewart, N Murphy, A Walker, A McGuire, and J J V McMurray. “Cost of an emerging epidemic: An economic analysis of atrial fibrillation in the UK”. *Heart* 90.3 (2004), pp. 286–292.
- [7] Johannes Brachmann et al. “Atrial fibrillation burden and clinical outcomes in heart failure: The CASTLE-AF trial”. *Clinical Electrophysiology* 7.5 (2021), pp. 594–603.
- [8] Luigi Di Biase et al. “Ablation versus amiodarone for treatment of persistent atrial fibrillation in patients with congestive heart failure and an implanted device: Results from the AATAC multicenter randomized trial”. *Circulation* 133.17 (2016), pp. 1637–1644.

- [9] Karl-Heinz Kuck et al. “Cryoballoon or radiofrequency ablation for paroxysmal atrial fibrillation”. *New England Journal of Medicine* 374.23 (2016), pp. 2235–2245.
- [10] Daniel Scherr et al. “Five-year outcome of catheter ablation of persistent atrial fibrillation using termination of atrial fibrillation as a procedural endpoint”. *Circulation: Arrhythmia and Electrophysiology* 8.1 (2015), pp. 18–24.
- [11] Jung Myung Lee et al. “The electrical isolation of the left atrial posterior wall in catheter ablation of persistent atrial fibrillation”. *Clinical Electrophysiology* 5.11 (2019), pp. 1253–1261.
- [12] Atul Verma et al. “Approaches to catheter ablation for persistent atrial fibrillation”. *New England Journal of Medicine* 372.19 (2015), pp. 1812–1822.
- [13] Julia Vogler et al. “Pulmonary vein isolation versus defragmentation: The CHASE-AF clinical trial”. *Journal of the American College of Cardiology* 66.24 (2015), pp. 2743–2752.
- [14] Sanjiv M. Narayan et al. “Treatment of atrial fibrillation by the ablation of localized sources: CONFIRM (Conventional Ablation for Atrial Fibrillation With or Without Focal Impulse and Rotor Modulation) trial”. *Journal of the American College of Cardiology* 60.7 (2012), pp. 628–636.
- [15] Kelvin CK Wong et al. “No benefit of complex fractionated atrial electrogram ablation in addition to circumferential pulmonary vein ablation and linear ablation: Benefit of complex ablation study”. *Circulation: Arrhythmia and Electrophysiology* 8.6 (2015), pp. 1316–1324.
- [16] Fu Siong Ng, Balvinder S. Handa, Xinyang Li, and Nicholas S. Peters. “Toward mechanism-directed electrophenotype-based treatments for atrial fibrillation”. *Frontiers in Physiology* 11 (2020).
- [17] Ozan Özgül et al. “High-density and high coverage composite mapping of repetitive atrial activation patterns”. *Computers in Biology and Medicine* 159 (2023), p. 106920.
- [18] Eric Buch et al. “Long-term clinical outcomes of focal impulse and rotor modulation for treatment of atrial fibrillation: A multicenter experience”. *Heart Rhythm* 13.3 (2016), pp. 636–641.
- [19] Yoram Rudy Ph.D. and John E. Burnes M.S. “Noninvasive electrocardiographic imaging”. *Annals of Noninvasive Electrocardiology* 4.3 (1999), pp. 340–359.
- [20] Caroline H. Roney et al. “Spatial resolution requirements for accurate identification of drivers of atrial fibrillation”. *Circulation: Arrhythmia and Electrophysiology* 10.5 (2017), e004899.
- [21] Renato E. Mirollo and Steven H. Strogatz. “Synchronization of pulse-coupled biological oscillators”. *SIAM Journal on Applied Mathematics* 50.6 (1990), pp. 1645–1662.
- [22] Miguel Antonio Aon, Sonia Cortassa, and Brian O’Rourke. “The fundamental organization of cardiac mitochondria as a network of coupled oscillators”. *Biophysical Journal* 91.11 (2006), pp. 4317–4327.
- [23] Zhilin Qu, Gang Hu, Alan Garfinkel, and James N. Weiss. “Nonlinear and stochastic dynamics in the heart”. *Physics Reports* 543.2 (2014), pp. 61–162.
- [24] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. “The graph neural network model”. *IEEE Transactions on Neural Networks* 20.1 (2008), pp. 61–80.
- [25] Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. “Geometric deep learning: Going beyond Euclidean data”. *IEEE Signal Processing Magazine* 34.4 (2017), pp. 18–42.
- [26] Davide Bacciu, Federico Errica, Alessio Micheli, and Marco Podda. “A gentle introduction to deep learning for graphs”. *Neural Networks* 129 (2020), pp. 203–221.
- [27] Youngjoo Seo, Michaël Defferrard, Pierre Vandergheynst, and Xavier Bresson. “Structured sequence modeling with graph convolutional recurrent networks”. *Proceedings of the 25th International Conference on Neural Information Processing*. Springer, 2018, pp. 362–373.
- [28] Andrea Cini, Ivan Marisca, and Cesare Alippi. “Filling the g_ap_s: Multivariate time series imputation by graph neural networks”. *Proceedings of the International Conference on Learning Representations* (2022).
- [29] Yang Li and Tao Yang. “Word embedding for understanding natural language: A survey”. *Guide to Big Data Applications* (2018), pp. 83–104.
- [30] Natasja MS De Groot et al. “Critical appraisal of technologies to assess electrical activity during atrial fibrillation: A position paper from the European heart rhythm association and European society of cardiology working group on eCardiology in collaboration with the heart rhythm society, Asia pacific heart rhythm society, Latin American heart rhythm society and computing in cardiology”. *EP Europace* 24.2 (2022), pp. 313–330.
- [31] Rui Shi et al. “Validation of dipole density mapping during atrial fibrillation and sinus rhythm in human left atrium”. *Clinical Electrophysiology* 6.2 (2020), pp. 171–181.
- [32] Xinyang Li et al. “Standardised framework for quantitative analysis of fibrillation dynamics”. *Scientific Reports* 9.1 (2019), p. 16671.
- [33] Laurens Van der Maaten and Geoffrey Hinton. “Visualizing data using t-SNE.” *Journal of Machine Learning Research* 9.11 (2008).
- [34] Norbert Marwan, M Carmen Romano, Marco Thiel, and Jurgen Kurths. “Recurrence plots for the analysis of complex systems”. *Physics Reports* 438.5-6 (2007), pp. 237–329.
- [35] Flavio Fenton and Alain Karma. “Vortex dynamics in three-dimensional continuous myocardium with fiber rotation: Filament instability and fibrillation”. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 8.1 (1998), pp. 20–47.

- [36] R. H. Clayton et al. “Models of cardiac tissue electrophysiology: Progress, challenges and open questions”. *Progress in Biophysics and Molecular Biology* 104.1-3 (2011), pp. 22–48.
- [37] Andrea Cini, Danilo Mandic, and Cesare Alippi. “Graph-based time series clustering for end-to-end hierarchical forecasting”. *Proceedings of the 41st International Conference on Machine Learning* (2024).
- [38] Ivan Marisca, Andrea Cini, and Cesare Alippi. “Learning to reconstruct missing data from spatiotemporal graphs with sparse observations”. *Proceedings of the 36th International Conference on Neural Information Processing Systems* (2022), pp. 32069–32082.
- [39] Antonio Ortega, Pascal Frossard, Jelena Kovačević, José M. F. Moura, and Pierre Vandergheynst. “Graph signal processing: Overview, challenges, and applications”. *Proceedings of the IEEE* 106.5 (2018), pp. 808–828.
- [40] Ljubiša Stanković et al. “Data analytics on graphs. Part II: Signals on graphs”. *Foundations and Trends® in Machine Learning* 13.2-3 (2020), pp. 158–331. ISSN: 1935-8237.
- [41] Ljubiša Stanković et al. “Data analytics on graphs. Part III: Machine learning on graphs, from graph topology to applications”. *Foundations and Trends® in Machine Learning* 13.4 (2020), pp. 332–530. ISSN: 1935-8237.
- [42] Thomas N. Kipf and Max Welling. “Semi-supervised classification with graph convolutional networks”. *Proceedings of the International Conference on Learning Representations* (2017).
- [43] Ziwei Zhang, Peng Cui, and Wenwu Zhu. “Deep learning on graphs: A survey”. *IEEE Transactions on Knowledge and Data Engineering* 34.1 (2020), pp. 249–270.
- [44] Andrea Cini, Ivan Marisca, Daniele Zambon, and Cesare Alippi. “Graph deep learning for time series forecasting”. *arXiv preprint arXiv:2310.15978* (2023).
- [45] Ming Jin et al. “A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024).
- [46] Jianfei Gao and Bruno Ribeiro. “On the equivalence between temporal and static equivariant graph representations”. *Proceedings of the 39th International Conference on Machine Learning* (2022), pp. 7052–7076.
- [47] Andrea Cini, Ivan Marisca, Filippo Maria Bianchi, and Cesare Alippi. “Scalable spatiotemporal graph neural networks”. *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. 2023. Chap. 6, pp. 7218–7226.
- [48] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. “Diffusion convolutional recurrent neural network: Data-driven traffic forecasting”. *Proceedings of the International Conference on Learning Representations* (2018).
- [49] Weiwei Jiang and Jiayun Luo. “Graph neural network for traffic forecasting: A survey”. *Expert Systems with Applications* 207 (2022), p. 117921.
- [50] Remi Lam et al. “Learning skillful medium-range global weather forecasting”. *Science* 382.6677 (2023), pp. 1416–1421.
- [51] Andrea Cini, Ivan Marisca, Daniele Zambon, and Cesare Alippi. “Taming local effects in graph-based spatiotemporal forecasting”. *Proceedings of the 37th International Conference on Neural Information Processing Systems* 36 (2024).
- [52] Xueyan Yin, Feifan Li, Yanming Shen, Heng Qi, and Baocai Yin. “NodeTrans: A graph transfer learning approach for traffic prediction”. *arXiv preprint arXiv:2207.01301* (2022).
- [53] Ruslan Salakhutdinov and Andriy Mnih. “Bayesian probabilistic matrix factorization using Markov chain Monte Carlo”. *Proceedings of the 25th International Conference on Machine Learning*, 2008, pp. 880–887.
- [54] Daniel Lee and H Sebastian Seung. “Algorithms for non-negative matrix factorization”. *Proceedings of the 13th International Conference on Neural Information Processing Systems* 13 (2000).
- [55] Deng Cai, Xiaofei He, Jiawei Han, and Thomas S Huang. “Graph regularized nonnegative matrix factorization for data representation”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.8 (2010), pp. 1548–1560.
- [56] Hsiang-Fu Yu, Nikhil Rao, and Inderjit S Dhillon. “Temporal regularized matrix factorization for high-dimensional time series prediction”. *Proceedings of the 30th International Conference on Neural Information Processing Systems* 29 (2016).
- [57] Wei Cao et al. “BRITS: Bidirectional recurrent imputation for time series”. *Proceedings of the 32nd International Conference on Neural Information Processing Systems* 31 (2018).
- [58] Jinsung Yoon, Daniel Jarrett, and Mihaela Van der Schaar. “Time-series generative adversarial networks”. *Proceedings of the 33rd International Conference on Neural Information Processing Systems* 32 (2019).
- [59] Yukai Liu, Rose Yu, Stephan Zheng, Eric Zhan, and Yisong Yue. “NAOMI: Non-autoregressive multiresolution sequence imputation”. *Proceedings of the 33rd International Conference on Neural Information Processing Systems* 32 (2019).
- [60] Wenjie Du, David Côté, and Yan Liu. “SAITS: Self-attention-based imputation for time series”. *Expert Systems with Applications* 219 (2023), p. 119619.
- [61] Yusuke Tashiro, Jiaming Song, Yang Song, and Stefano Ermon. “CSDI: Conditional score-based diffusion models for probabilistic time series imputation”. *Proceedings of the 35th International Conference on Neural Information Processing Systems* 34 (2021), pp. 24804–24816.
- [62] J Alcaraz and N Strodthoff. “Diffusion-based time series imputation and forecasting with structured state space models”. *Transactions on Machine Learning Research* (2023).
- [63] Alexander Jenkins, Zehua Chen, Fu Siong Ng, and Danilo Mandic. “Improving diffusion models for ECG imputation with an augmented template prior”. *arXiv preprint arXiv:2310.15742* (2023).
- [64] Pablo Montero-Manso and Rob J Hyndman. “Principles and algorithms for forecasting groups of time series: Locality and globality”. *International Journal of Forecasting* 37.4 (2021), pp. 1632–1653.

- [65] Xavier Bresson and Thomas Laurent. “Residual gated graph convnets”. *arXiv preprint arXiv:1711.07553* (2017).
- [66] James Atwood and Don Towsley. “Diffusion-convolutional neural networks”. *Proceedings of the 30th International Conference on Neural Information Processing Systems* 29 (2016).
- [67] Roger Koenker and Kevin F Hallock. “Quantile regression”. *Journal of Economic Perspectives* 15.4 (2001), pp. 143–156.
- [68] Mikhail Belkin and Partha Niyogi. “Laplacian eigenmaps for dimensionality reduction and data representation”. *Neural Computation* 15.6 (2003), pp. 1373–1396.
- [69] Paul J Besl and Neil D McKay. “Method for registration of 3-D shapes”. *Proceedings of Sensor fusion IV: Control Paradigms and Data Structures* 1611 (1992), pp. 586–606.

Supplementary materials

S.1 Background and related work	14
S.2 FIBMAP	14
S.2.1 Inputs	15
S.2.2 Architecture	15
S.2.2.1 The spatiotemporal encoder	15
S.2.2.2 The decoder	17
S.2.3 Optimisation	17
S.3 Data and implementation details	17
S.3.1 Dataset	17
S.3.2 Simulating sequential contact mapping	18
S.3.3 Training FIBMAP	19
S.3.4 Fine-tuning FIBMAP	19
S.3.5 Performance metrics	20
S.3.6 Baseline models	20
S.3.7 Imputation mapping from EnSite Precision HD Grid recordings	21

S.1 Background and related work

The field of graph signal processing [39, 40] has emerged in the last decade to generalise digital signal processing methods, such as convolutional filters, to the graph domain. A graph, $\mathcal{G}(\mathcal{V}, \mathcal{E})$, is defined as a set of N nodes, $v_i \in \mathcal{V}$ for $i = 1, \dots, N$, and a set of \mathcal{E} edges denoting the pairwise connections between them, $e_{ij} = (v_i, v_j) \in \mathcal{E}$, for $i = 1, \dots, N$ and $j = 1, \dots, N$. Graph signal processing focuses on analysing signals on graphs, where a multivariate signal on a graph is defined as $\mathbf{X} \in \mathbb{R}^{N \times c}$, which assigns a real-valued signal of c channels to each node, $\mathbf{X} : \mathcal{V} \mapsto \mathbb{R}^c$.

Building on graph signal processing methods, graph deep learning [25, 41] has generalised successful deep learning architectures, such as convolutional neural networks, to the graph domain [42, 43]. Spatio-temporal graph neural networks (ST-GNNs) refer to the class of deep learning architectures designed to analyse time-varying graph signals [44, 45]. ST-GNNs can be categorised into time-then-space [46, 47] or time-and-space models [27, 48, 38], which denote separate or joint processing of the space and time dimensions, respectively. A notable example of a time-and-space ST-GNN is the GRNN, which replaces the fully connected layers in a recurrent neural network (RNN) with graph convolutions [27, 48]. ST-GNNs have achieved state-of-the-art performance in tasks such as forecasting [49, 50, 37] and missing data imputation [38, 28].

ST-GNNs are inherently global models, sharing parameters across space and assuming a stationary process over time. Global ST-GNNs can be used for zero-shot transfer and inductive learning on unseen graphs, however, they might fail to account for spatial variations in dynamics. Node embeddings have been recently introduced to learn these local effects in ST-GNNs [51]. Whilst hybrid global-local models often outperform global architectures, recent research has focussed on improving their performance in an inductive setting using transfer learning [51, 52].

Other imputation methods have been applied to time series, ranging from MF methods [53, 54] and their counterparts with graph and temporal regularisation [55, 56], to deep learning techniques employing RNNs [57], generative adversarial networks [58, 59], transformers [60], or most recently denoising diffusion probabilistic models [61, 62, 63].

S.2 FIBMAP

FIBMAP aims to reconstruct the dynamics of AF from partial measurements collected during sequential contact mapping using a GRNN. This task is formulated using a spatiotemporal time series imputation framework, which models the propagation of electrical signals across the atrium via spatiotemporal message passing on a graph.

To become a feasible clinical solution, the design of FIBMAP solves several technical challenges of the task: 1) how to effectively propagate information from highly sparse sets of observations where up to 10% of the atrium is observed per time with sequential contact mapping; 2) how to manage possibly unknown parameters which could affect the dynamics, such as the spatial properties of the tissue and patient-specific covariates; 3) how to efficiently train and use a single model within the duration of the clinical procedure, whilst balancing the trade-off between personalised reconstruction and generalising across a spectrum of AF dynamics; and 4) how to quantify the uncertainty of the estimated reconstruction — an essential factor for ensuring a useful and trustworthy clinical tool in practice. In this section, we describe the framework behind FIBMAP, our solution for achieving accurate reconstruction of AF dynamics while considering the technical challenges detailed above.

S.2.1 Inputs

A graph adjacency matrix for each patient p , $\mathbf{A}^{(p)} \in \mathbb{R}^{N \times N}$, is created by discretising the atrial surface into N nodes and triangulating the surface to create edges. The adjacency matrix, $\mathbf{A}^{(p)}$, is constant over time. A sequence of graphs is formed as a tuple, $\mathcal{G}_{t:T}^{(p)} = \langle \mathbf{X}_{t:T}^{(p)}, \mathbf{M}_{t:T}^{(p)}, \mathbf{A}^{(p)} \rangle$, whereby partial electrical signals, $\mathbf{X}_{t:T}^{(p)} \in \mathbb{R}^{N \times 1}$, are observed via sequential contact mapping, and the observation mask, $\mathbf{M}_{t:T}^{(p)} \in \{0, 1\}^{N \times 1}$, specifies the indices of the valid measurements.

To manage the patient’s latent spatial and global parameters which could affect the dynamics of AF, trainable embeddings are used as an additional input. Specifically, trainable node embeddings, $\mathbf{V}^{(p)} \in \mathbb{R}^{N \times q}$, are used to learn the latent spatial properties of the patient’s tissue, and trainable patient embeddings, $\mathbf{g}^{(p)} \in \mathbb{R}^r$, will be used to learn the latent global properties of the patient. In this way, the sequence of graph data for each patient can be expressed as a tuple, $\mathcal{G}_{t:L}^{(p)} = \langle \mathbf{X}_{t:L}^{(p)}, \mathbf{M}_{t:L}^{(p)}, \mathbf{A}^{(p)}, \mathbf{V}^{(p)}, \mathbf{g}^{(p)} \rangle$. For efficient processing, input sequences are formed by splitting the sequence of length, L , into windows of length, W , processed with a stride, M .

S.2.2 Architecture

FIBMAP is instantiated as a bidirectional gated-GRNN, a non-linear state-space model designed to propagate information from valid observations across space and time. Our solution is composed of two spatiotemporal encoder blocks and a decoder. The spatiotemporal encoders operate in two different directions, processing the sequence in both a forward (*fwd*) and backward (*bwd*) direction, respectively. FIBMAP’s architecture extends the framework proposed in [28] by introducing local and global embedding parameters to address the AF mapping problem. These modifications enable generalisation across patients and efficient transfer to new patients via fine-tuning.

To balance the trade-off between personalised reconstruction and generalisation across a spectrum of AF dynamics with a single model, FIBMAP is designed to perform personalised reconstruction by providing patient-specific parameters, $\mathbf{V}^{(p)}$ and $\mathbf{g}^{(p)}$, as additional input to the encoder and decoder components of the architecture. All other model parameters are shared across patients to learn the common aspects of the dynamics (such as the physics of the problem). We begin by defining the spatiotemporal encoder block of the architecture, before explaining the decoder.

S.2.2.1 The spatiotemporal encoder

First, the observed data at time t is encoded as

$$\mathbf{H}_t^0 = \text{MLP}_{\text{enc}}([\mathbf{X}_t^{(p)} \odot \mathbf{M}_t^{(p)} \parallel \mathbf{M}_t^{(p)} \parallel \mathbf{G}^{(p)} \parallel \mathbf{V}^{(p)}]), \quad (1)$$

where the symbols \odot and \parallel denote the Hadamard product and concatenation operator, respectively, and $\mathbf{G}^{(p)} \in \mathbb{R}^{N \times r}$ is a matrix where copies of $\mathbf{g}^{(p)}$ are stacked across nodes. The encoding step constructs a representation of the observed and missing values alongside the patient-specific parameters in a common embedding space, $\mathbf{H}_t^0 \in \mathbb{R}^{N \times d}$. Next, the embedded data is processed sequentially using a gated-GRNN, where the k -th layer is given by

$$\mathbf{Z}_t^k = \mathbf{H}_{t-1}^k, \quad (2a)$$

$$\mathbf{R}_t^k = \sigma(\text{MP}_r^k([\mathbf{Z}_t^k \parallel \mathbf{H}_{t-1}^k], \mathcal{E})), \quad (2b)$$

$$\mathbf{U}_t^k = \sigma(\text{MP}_u^k([\mathbf{Z}_t^k \parallel \mathbf{H}_{t-1}^k], \mathcal{E})), \quad (2c)$$

$$\mathbf{C}_t^k = \tanh(\text{MP}_c^k([\mathbf{Z}_t^k \parallel \mathbf{R}_t^k \odot \mathbf{H}_{t-1}^k], \mathcal{E})), \quad (2d)$$

$$\mathbf{H}_t^k = \mathbf{U}_t^k \odot \mathbf{H}_{t-1}^k + (1 - \mathbf{U}_t^k) \odot \mathbf{C}_t^k. \quad (2e)$$

$\text{MP}_r^k(\cdot)$, $\text{MP}_u^k(\cdot)$ and $\text{MP}_c^k(\cdot)$ denote the message passing neural network (MPNN) layers for the reset, update, and candidate gates, respectively, and the activation functions $\sigma(\cdot)$ and $\tanh(\cdot)$ denote the sigmoid and hyperbolic tangents.

The hidden state of the gated-GRNN at each layer k is initialised as a linear function of the node embeddings

$$\mathbf{H}_{t=0}^k = \mathbf{V}^{(p)} \mathbf{W}_{\text{init}}^k + \mathbf{b}_{\text{init}}^k, \quad (3)$$

where $\mathbf{W}_{\text{init}}^k \in \mathbb{R}^{q \times d}$ and $\mathbf{b}_{\text{init}}^k \in \mathbb{R}^d$, which takes the different characteristics of each time series into account when initialising the state, thus reducing the need to rely on a long observation history [64].

Message passing layers The MPNN layers for each gate in the GRNN, denoted as MP, compute the hidden state of the i -th node as

$$\mathbf{h}_{t,i}^{k+1} = \gamma^k \left(\mathbf{h}_{t,i}^k, \sum_{j \in \mathcal{N}(i)} \rho^k(\mathbf{o}_i^k, \mathbf{o}_j^k, e_{ij}) \right), \quad (4)$$

where $\mathbf{o}_i^k \in \mathbb{R}^{2d}$ represents the input of the gate at layer k , $\mathcal{N}(i)$ denotes the set of nodes connected to i , $\rho^k(\cdot)$ is a message function, $\gamma^k(\cdot)$ is an update function, and the summation serves as a permutation invariant aggregation over neighbouring nodes. Specifically, the message function $\rho^k(\cdot)$ is implemented as

$$\mathbf{m}_{ij}^k = \text{MLP}_{\text{msg}}^k([\mathbf{o}_i^k \| \mathbf{o}_j^k]), \quad (5a)$$

$$\alpha_{ij}^k = \sigma(\mathbf{W}_{\text{msg}-\alpha} \mathbf{m}_{ij}^k), \quad (5b)$$

$$\mathbf{m}_{ij}^k = \alpha_{ij}^k \mathbf{m}_{ij}^k, \quad (5c)$$

where the messages along non-zero edges e_{ij} are weighted according to an inferred scalar value α_{ij}^k , which resides on the interval $[0, 1]$. The scalar value α_{ij}^k allows for anisotropic message passing, which assists in learning latent edge attributes such as the coupling strength between nodes. Messages are then aggregated at node i using the summation, $\mathbf{m}_i^k = \sum_{j \in \mathcal{N}(i)} \mathbf{m}_{ij}^k$. Finally, the update function $\gamma^k(\cdot)$ is implemented with a residual skip connection as

$$\mathbf{u}_i^k = \text{MLP}_{\text{update}}^k([\mathbf{h}_i^k \| \mathbf{m}_i^k]), \quad (6a)$$

$$\mathbf{h}_{t,i}^{k+1} = \mathbf{u}_i^k + \mathbf{W}_{\text{skip}} \mathbf{o}_i^k, \quad (6b)$$

where $\mathbf{W}_{\text{skip}} \in \mathbb{R}^{d \times 2d}$ and $\mathbf{h}_i^{k+1} \in \mathbb{R}^d$. Note that the parameters of the MPNN layers, MP, are defined separately for each gate and layer. This message-passing mechanism loosely resembles the gated graph convolution introduced in [65].

Iterative imputations To learn effective state space representations while processing the data sequentially, missing values must be accounted for. To do this, first and second-stage imputations as in [28] are performed iteratively within the spatiotemporal encoder block. The first stage imputation performs one-step-ahead prediction via a linear readout as

$$\hat{\mathbf{X}}_{t+1}^{(1)} = \mathbf{H}_t^K \mathbf{W}_{\text{stage-1}} + \mathbf{b}_{\text{stage-1}}, \quad (7)$$

which is used in place for the missing values. The second stage imputation acts as a type of regularisation, estimating the value at node i using the previous hidden states, as well as the masks and first-stage imputations at the neighbouring nodes. The second stage imputation involves first computing an intermediate representation of each node, given by

$$\mathbf{s}_{t+1,i} = \gamma_{\text{stage-2}} \left(\mathbf{h}_{t,i}^K, \sum_{j \in \mathcal{N}(i)} \rho_{\text{stage-2}}([\hat{\mathbf{x}}_{t+1,j}^{(1)} \| \mathbf{h}_{t,j}^K \| \mathbf{m}_{t+1,j}], e_{ij}) \right), \quad (8)$$

where $\mathbf{s}_{t+1,i} \in \mathbb{R}^d$ and the message passing is implemented using a diffusion graph convolution [66]. Next, a linear readout layer performs the second stage imputation as

$$\hat{\mathbf{X}}_{t+1}^{(2)} = [\mathbf{s}_{t+1} \| \mathbf{H}_t^K] \mathbf{W}_{\text{stage-2}} + \mathbf{b}_{\text{stage-2}}, \quad (9)$$

which is used again in place of the missing values in \mathbf{X}_{t+1} .

The process detailed so far in the spatiotemporal encoding block is repeated for \mathbf{X}_{t+1} , to learn representations \mathbf{S}_{t+2} and \mathbf{H}_{t+1}^K , and so forth, until a set of representations $\mathbf{S}_{t:t+W}$ and $\mathbf{H}_{t:t+W}^K$, and imputations $\hat{\mathbf{X}}_{t:t+W}^{(1)}$ and $\hat{\mathbf{X}}_{t:t+W}^{(2)}$ are calculated for the whole window length W . To summarise the spatiotemporal encoding block detailed in this section, we use the following shorthand notation

$$\langle \mathbf{S}_{t:t+W}, \mathbf{H}_{t:t+W}^K, \hat{\mathbf{X}}_{t:t+W}^{(1)}, \hat{\mathbf{X}}_{t:t+W}^{(2)} \rangle = \text{ST-Encoder}(\mathbf{X}_{t:t+W}^{(p)}, \mathbf{M}_{t:t+W}^{(p)}, \mathbf{g}^{(p)}, \mathbf{V}^{(p)}). \quad (10)$$

S.2.2.2 The decoder

The sequences are encoded in both the forward and backward directions, to form

$$\langle \mathbf{S}_{\text{fwd}}, \mathbf{H}_{\text{fwd}}^K, \widehat{\mathbf{X}}_{\text{fwd}}^{(1)}, \widehat{\mathbf{X}}_{\text{fwd}}^{(2)} \rangle = \text{ST-Encoder}(\mathbf{X}_{t:t+W}^{(p)}, \mathbf{M}_{t:t+W}^{(p)}, \mathbf{g}^{(p)}, \mathbf{V}^{(p)}), \quad (11)$$

and

$$\langle \mathbf{S}_{\text{bwd}}, \mathbf{H}_{\text{bwd}}^K, \widehat{\mathbf{X}}_{\text{bwd}}^{(1)}, \widehat{\mathbf{X}}_{\text{bwd}}^{(2)} \rangle = \text{ST-Encoder}(\mathbf{X}_{t+W:t}^{(p)}, \mathbf{M}_{t+W:t}^{(p)}, \mathbf{g}^{(p)}, \mathbf{V}^{(p)}), \quad (12)$$

respectively, where $t + W : t$ denotes the time reversed/backward sequence. The next stage is to decode the hidden state representations from the encodings in both directions, to perform a final imputation at time t' . This is done with the following function

$$\widehat{\mathbf{Y}}_{t'} = \text{MLP}_{\text{dec}}([\mathbf{S}_{\text{fwd}} \leq t' \| \mathbf{H}_{\text{fwd}}^K \leq t' \| \mathbf{S}_{\text{bwd}} \geq t' \| \mathbf{H}_{\text{bwd}}^K \geq t' \| \mathbf{M}_{t'}^p \| \mathbf{V}^{(p)} \| \mathbf{g}^{(p)}]), \quad (13)$$

where the notation $\text{fwd} \leq t'$ refers to an index of the hidden states at time t' (which includes information from times $\leq t'$) and similarly, $\text{bwd} \geq t'$ refers to index at time t' (which includes information from times $\geq t'$). The decoder takes a form similar to that in [28] except patient-specific parameters are also provided to facilitate personalisation. We denote the final imputed values for the entire window length as $\widehat{\mathbf{Y}}_{t:t+W}$.

S.2.3 Optimisation

To quantify the uncertainty of the reconstructed dynamics, FibMAP is formulated as a quantile regression [67]. In general, a quantile regression aims to estimate the conditional quantile function, $Q_Y(\tau|\mathbf{B})$, which represents the τ -th quantile of the response variable, Y , given the predictor variables, \mathbf{B} . This is given by $Q_Y(\tau|\mathbf{B}) = \inf\{y \in \mathbb{R} | P(Y \leq y|\mathbf{B}) \geq \tau\}$, where $P(Y \leq y|\mathbf{B})$ is the conditional cumulative distribution function of the response variable, Y , given the predictors \mathbf{B} , and \inf represents the infimum of the set.

In this work, conditional quantile functions for $\tau \in \mathcal{C}$, where $\mathcal{C} = \{\tau_1, \tau_2, \dots, \tau_{|\mathcal{C}|}\}$, are predicted for each of the imputed values. This is done by computing the following masked quantile/pinball loss function

$$\mathcal{L}(\widehat{\mathbf{Y}}_{t:t+T}, \widetilde{\mathbf{Y}}_{t:t+T}, \bar{\mathbf{M}}_{t:t+T}) = \frac{\sum_{h=t}^{t+T} \sum_{i=1}^N \bar{m}_{t,i} \mathcal{L}_{t,i}(\widehat{\mathbf{y}}_{t,i}, y_{t,i})}{\sum_{h=t}^{t+T} \sum_{i=1}^N \bar{m}_{t,i}}, \quad (14)$$

where $\widehat{\mathbf{Y}}_{t:t+T} \in \mathbb{R}^{N \times |\mathcal{C}|}$, $\mathbf{Y}_t \in \mathbb{R}^{N \times 1}$, and $\bar{\mathbf{M}}_t \in \{0, 1\}^{N \times 1}$, represent the predicted values, target values, and the evaluation mask, respectively, at time t . The function $\mathcal{L}_{t,i}$ computes the average pinball loss computed at time t and node i , which is given by

$$\mathcal{L}_{t,i}(\widehat{\mathbf{y}}_{t,i}, y_{t,i}) = \frac{1}{|\mathcal{C}|} \sum_{c=1}^{|\mathcal{C}|} (y_{t,i} - \widehat{\mathbf{y}}_{t,i}[c]) (\tau_c - \mathbb{1}\{y_{t,i} - \widehat{\mathbf{y}}_{t,i}[c] < 0\}), \quad (15)$$

where c refers to the channel of $\widehat{\mathbf{y}}_{t,i}$ (prediction vector at time t and node i) which contains the prediction of the c -th conditional quantile, $Q_y(\tau_c|\dots)$ at time t and node i , and $\mathbb{1}\{y_{t,i} - \widehat{\mathbf{y}}_{t,i}[c] < 0\}$ represents the indicator function, which is equal to 1 if $y_{t,i} - \widehat{\mathbf{y}}_{t,i}[c] < 0$, else it is equal to 0. In practice, we modify the decoder to predict a value for each quantile (rather than a single value) and compute the average pinball loss from each predicted quantile value as in (15).

For FibMAP, the following loss function is minimised

$$\mathcal{L}_{\text{loss}} = \mathcal{L}(\widehat{\mathbf{Y}}_{t:t+T}, \mathbf{X}_{t:t+T}, \bar{\mathbf{M}}_{t:t+T}) \quad (16)$$

$$+ \mathcal{L}(\widehat{\mathbf{X}}_{t:t+T}^{(1),\text{fwd}}, \mathbf{X}_{t:t+T}, \bar{\mathbf{M}}_{t:t+T}) + \mathcal{L}(\widehat{\mathbf{X}}_{t:t+T}^{(2),\text{fwd}}, \mathbf{X}_{t:t+T}, \bar{\mathbf{M}}_{t:t+T}) \quad (17)$$

$$+ \mathcal{L}(\widehat{\mathbf{X}}_{t:t+T}^{(1),\text{bwd}}, \mathbf{X}_{t:t+T}, \bar{\mathbf{M}}_{t:t+T}) + \mathcal{L}(\widehat{\mathbf{X}}_{t:t+T}^{(1),\text{bwd}}, \mathbf{X}_{t:t+T}, \bar{\mathbf{M}}_{t:t+T}), \quad (18)$$

where \mathcal{L} is given in (14) and each component of the loss is specific to a different imputation stage and processing direction. The specification of the evaluation mask, $\bar{\mathbf{M}}$, differs depending on training and fine-tuning as detailed in the next sections.

S.3 Data and implementation details

S.3.1 Dataset

Our dataset comprises recordings from 51 patients with persistent AF, obtained using the AcQMap system. The system uses non-contact electrodes to sense intra-cavitary electrograms, from which dipole density measurements (charge

Table S.T1: Patient demographics and clinical characteristics. Values are mean \pm SD or n (%). The asterisk denotes the presence of missing values; the number missing is shown as n^* .

Total ($N = 51$)	
Demographics	
Mean age (years)	64 \pm 11
Male sex	35 (69%)
Mean BMI	29 \pm 5
Comorbidities	
Hypertension	20 (39%)
Diabetes mellitus	5 (10%)
Heart failure	17 (33%)
CHA ₂ DS ₂ -VASc score >2	14 (27%)
Medications	
Beta-blockers	33* (87%, $n^* = 13$)
Amiodarone	11* (29%, $n^* = 13$)
Anticoagulants	38* (100%, $n^* = 13$)
Statins	16* (42%, $n^* = 13$)

density in Coulombs/cm) are inversely derived across the entire atria [30, 31]. Each recording samples approximately 3500 nodes at 3000 Hz for 5-20 seconds. All recordings were obtained prior to pulmonary vein isolation at two United Kingdom centres between 2016 and 2023. The patient cohort (mean age 64 ± 11 years, 69% male) had standard cardiovascular risk factors and were on guideline-directed medical therapy (see Table S.T1 for more details).

For pre-processing, these signals are first resampled spatially to a resolution consisting of 500 nodes. This is done by k -means clustering of the 3D node coordinates into $k = 500$ non-overlapping clusters, where signals are resampled using the mean average signal within each cluster for each time point. Temporal resampling to 70 Hz is conducted through a combination of low pass filtering and down sampling. Low-pass filtering is applied to each time series at 70 Hz to prevent aliasing, and then the filtered signal is downsampled by reducing the sampling rate proportionally. Finally, the time series are normalised across space and time by applying min-max normalisation, where the minimum and maximum values are determined across all nodes and times, ensuring that the amplitude of the resampled signals is scaled consistently between 0 and 1.

The dataset was stratified to ensure a spectrum of AF dynamics within training (70%), validation (10%), and test (20%) sets. The Shannon entropy was first computed for the resampled time series at each node to do this. The cumulative distribution function (CDF) of Shannon entropy across the atrium was then created for each patient, serving as a measure to compare the similarity of the spatiotemporal dynamics between patients. The similarity between patients was assessed by computing Kolmogorov–Smirnov (KS) distance between CDFs for each pair of patients, where a smaller KS distance represents more similar entropy distributions/spatiotemporal dynamics. From this, groups of similar patients were identified by clustering a Laplacian eigenmap [68] using k -means, where the number of clusters was determined using the elbow method. This involved forming a weighted adjacency matrix by applying a radial basis function kernel to the reciprocal of the KS distances and thresholding. Finally, the training, validation, and test sets were formed by performing stratified sampling across these clusters, maintaining a consistent spectrum of AF dynamics within each set.

S.3.2 Simulating sequential contact mapping

A sequential contact mapping strategy is simulated from the non-contact recordings as a self-avoiding walk of the multipolar catheter (represented as a patch of observations) on the atrial surface; whereby the catheter surface area, dwell time, and spatial overlap, could be controlled. The self-avoiding walk is defined by first providing the catheter surface area as a fraction of the total area, where its reciprocal (1/area) is rounded up to the nearest integer to give the number of patches required to sample the whole atrium. Then, the atrium is split into disjoint patches by k -means clustering the 3D node coordinates, with k equal to the number of patches. Spatial overlap between patches is simulated by adding additional clusters between adjacent patches and sharing node sets in proportion to the overlap required.

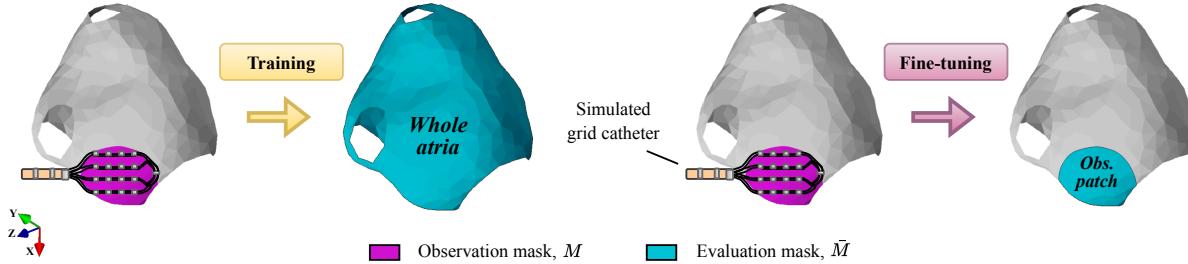


Figure S.F1: Definition of the observation and evaluation masks in FIBMAP training and fine-tuning. During training, FIBMAP is optimised to perform whole atria mapping using a whole atria evaluation mask to compute the loss function. During fine-tuning, only the patient-specific parameters of FIBMAP are optimised using an observed patch reconstruction loss.

A self-avoiding random walk ensures the path does not revisit the same region twice. This is simulated by sampling a patch randomly, then sampling the remaining patches without replacement with a probability proportional to the distance between the current and remaining patches. The resulting sequence of patches is converted to an observation mask, M , by assigning a unity value if nodes are within the observed patch, zero otherwise, and repeating these values such that each patch is observed for a duration equal to the specified dwell time. Except during the sensitivity analysis, the sequence of patches is repeated until the length of the available recording is met.

S.3.3 Training FIBMAP

The training procedure of FIBMAP leverages the whole atria ground truth signals of each patient to learn a robust function for reconstructing whole atrium dynamics from data collected in a sequential contact mapping fashion. To do this, a self-supervised approach is employed, wherein self-avoiding walks of the multipolar catheter are sampled at random to form the observation mask, M , and observed input data, \mathbf{X} , of each training sample, alongside a range of catheter surface areas (2.5 – 50% of the atrium), dwell times (0.2 - 4 seconds) and spatial overlaps (0 – 3 additional clusters between adjacent patches). At each training iteration, batches of size B are formed by collating samples from different patients, whereby B inputs $\mathbf{X}_{t:t+W}^{(p)}, \mathbf{M}_{t:t+W}^{(p)}, \mathbf{A}^{(p)}, \mathbf{V}^{(p)}, \mathbf{g}^{(p)}$ are collated for a temporal input window size, W , sampled from the original time series following a sliding window approach with unit stride.

Imputation is performed within the temporal window and a whole atria reconstruction loss is used to optimise FIBMAP during training, whereby (18) is optimised using an evaluation mask, $\bar{\mathbf{M}}_{t:t+W}$, with all nodes and times in the input window equal to unity (see Figure S.F1). This approach ensures a robust function for reconstructing whole atria maps that generalise across the distribution of multipolar catheter paths and parameters of sequential contact mapping such as dwell time.

A hyperparameter search is conducted for the parameters shown in Table S.T2, by first splitting the time series of each patient sequentially, with the first 85% of time steps being used for training, the second 5% for validation, and the final 10% for testing. For the validation and test sequences, a fixed self-avoiding walk of the multipolar catheter is used with a surface area of 10%, a dwell time of 1 second, and no spatial overlap. Each configuration in the hyperparameter search is conducted for 100 epochs, whereby the MAE across the remaining atria for the validation sequence is monitored with early stopping. All experiments were performed on a NVIDIA RTX A5000 graphics processing unit. The best-performing set of hyperparameters are chosen by computing the MAE loss across the remaining atria for the test sequence and selecting the minimum loss. The best-performing configuration for training FIBMAP was found to be $B = 16$, $W = 40$, $d = 64$, $q = 64$, $r = 16$, $K = 1$, and 1024 hidden layer neurons. This configuration was retrained for 500 epochs, where a learning rate of 0.0009 and a cosine scheduler were used. Training took a total of 31 hours. The result is a pre-trained FibMap model, which performs accurate and robust whole atria reconstruction from partial observations across a spectrum of dynamics found in patients and variations in multipolar mapping.

S.3.4 Fine-tuning FIBMAP

In clinical practice, only the patches of multipolar catheter measurements are available from sequential contact mapping. The rest of the atrium remains unobserved. To make FIBMAP a feasible clinical solution for imputation mapping, a fine-tuning procedure is introduced to enable accurate whole atria reconstruction on new and unseen patients when only partial observations of the dynamics are available.

Table S.T2: Hyperparameter values tested during FIBMAP training.

Hyperparameter	Range tested
Batch size, B	[16, 32, 64]
Input window length, W	[20, 30, 40, 50]
Hidden size, d	[64, 128, 256]
Node embedding size, q	[16, 32, 64]
Patient embedding size, r	[16, 32, 64]
Layers, K	[1, 2, 3]
Hidden layer neurons in MLP_{enc} , MLP_{dec}	[128, 256, 512, 1024]

Our fine-tuning procedure preserves essential knowledge for whole atria reconstruction acquired during training by fixing the parameters of the pre-trained FIBMAP model, while quickly personalising the model to new patients by optimising only the node and patient embedding parameters using an observed patch reconstruction loss (see Figure S.F1 for an illustration of the observation and evaluation masks used during fine-tuning, which are defined to be equal). The observed patch reconstruction loss is defined using this evaluation mask in (18). This restricts the optimisation during fine-tuning to only the patches of multipolar catheter measurements.

Our fine-tuning procedure was configured on the validation set, which also has ground truth whole atria signals available for patients. Validation of the fine-tuning procedures aims to evaluate the relationship between the observed patch and whole atria reconstruction losses. Again, the time series of each patient was split sequentially, with the first 85% of time steps being used for training, the second 5% for validation, and the final 10% for testing.

A random hyperparameter search was conducted across learning rates [0.0005, 0.005] and batch sizes [16, 32, 64, 128]. Each learning rate in the hyperparameter search was conducted for 100 epochs, whereby the MAE across the remaining atria for the validation sequence was monitored with early stopping. For the validation and test sequences, a fixed self-avoiding walk of the multipolar catheter was used with a surface area of 10%, a dwell time of 1 second, and no spatial overlap. The best-performing set of hyperparameters was chosen by computing the MAE across the remaining atria for the test sequence and selecting the minimum loss. The best-performing configuration for fine-tuning FIBMAP was found to have a learning rate of 0.005 and a batch size of 16.

Finally, the performance of FIBMAP imputation mapping on new and unseen patients, through our fine-tuning setup, was quantified on the test set patients. A sequential time series split was not used during testing, instead, all observed measurements were used. FIBMAP was fine-tuned for 100 epochs using the configuration found in validation, and the observed patch reconstruction loss was monitored to perform early stopping. The performance metrics, detailed next, were computed across the remaining atria to assess test set performance.

S.3.5 Performance metrics

Quantitative assessment of the reconstructed imputation maps was performed using several metrics: MAE, mean relative error (MRE) and MAPE. These metrics evaluate the fidelity of the reconstructed whole atria maps, $\hat{\mathbf{Y}}$, against the ground truth, \mathbf{X} , for the mask, $\hat{\mathbf{M}}$, which represents the logical binary complement of the observation mask, \mathbf{M} . The averaged performance metrics were computed via (14), where $\mathcal{L}_{t,i}$ was computed using each of our evaluation metrics. Metrics were computed for each model trained or fine-tuned across 5 different seeds. The average and standard deviation across these seeds were reported for each metric.

Additionally, the PS TPR was used to evaluate the accuracy of tracking PSs in the reconstructed phase maps compared to the ground truth. For each patient in the test set and each imputation model, PSs were manually labelled by two independent trained observers using a custom graphical user interface. Annotating each frame for each patient and model is labour-intensive, so only the central 70 frames (1 second) of each reconstructed phase map were annotated. Annotating PSs for each model seed is also infeasible, so the mean and standard deviation were computed from 1000 bootstrap samples. A PSs is considered detected if its location in the reconstructed map falls within a specified tolerance of 0.1 seconds and a 4-hop neighbourhood in space compared to its location in the ground truth map.

S.3.6 Baseline models

We introduce additional baseline models for the imputation mapping task:

1. Mean, which performs imputation using the node-level average;
2. MF with rank = 10;
3. Univariate RNN, which performs imputation based solely on the node-level signals;
4. Univariate bidirectional (Bi)-RNN.

Mean and MF baseline models are employed solely on the test set due to their transductive nature. Both the univariate RNN and Bi-RNN models were trained using MAE loss function and followed identical hyperparameter settings and training-test protocols as FIBMAP. Unlike FIBMAP, these models did not require fine-tuning due to their inductive nature.

S.3.7 Imputation mapping from EnSite Precision HD Grid recordings

From our test cohort, three AF patients had both EnSite Precision HD Grid and AcQMap recordings collected non-contemporaneously before ablation. For these patients, AcQMap recordings were 20 seconds in duration, while EnSite Precision HD Grid recordings were significantly longer (14, 19, and 20 minutes). EnSite Precision HD Grid data consists of sequential contact mapping recordings from 16 electrodes arranged in a grid array, along with the roving 3D coordinates of each electrode within the atrium.

To ensure compatibility with FIBMAP, the EnSite Precision HD Grid recordings were pre-processed. Sparse electrogram recordings were first mapped to a uniform discretisation of the atrial surface using a nearest-neighbour approach with a 3 mm radius to interpolate the signals between electrodes. An observation mask identified active recording periods, after which signals were normalised using the maximum peak-to-peak voltage. The data was then resampled spatially to a resolution of 500 nodes using k-means clustering and resampled temporally to 70 Hz using a combination of low-pass filtering and downsampling. A final min-max normalisation ensured all signals fell within 0 and 1. From these processed measurements, imputation maps were generated using our fine-tuning procedure, whereby the observed patch reconstruction loss was monitored to perform early stopping.

We developed a sliding window cross-correlation framework to compare FIBMAP reconstructions against AcQMap ‘ground truth’ recordings. While temporal alignment was not possible between the non-contemporaneous recordings, spatial alignment was achieved between AcQMap and imputation map by centring the vertices of the geometries around the origin, applying rigid registration using the iterative closest point algorithm [69], and projecting data between geometries using a $k = 5$ nearest-neighbours regression. Using sliding window lengths from 0.5 to 4.0 seconds and a constant stride of 0.1 seconds, we computed the Pearson correlations between Hilbert phases of processed signals across all nodes and times within window pairs. This generated a cross-correlation matrix characterising the spatiotemporal similarity between recordings, which were flattened and plotted as distributions for analysis.

To validate that FIBMAP captured patient-specific dynamics, we performed three types of correlations: intra-patient comparisons between AcQMap and FIBMAP from the same patient; inter-patient comparisons between AcQMap and FIBMAP from different patients; and random baseline intra-patient comparisons between AcQMap and a spatiotemporally shuffled imputation map. For each patient, the 99th percentile of the intra, inter and shuffled distributions were plotted and statistical significance was assessed by computing the confidence intervals via bootstrapping ($n = 1000$ rounds of 10000 resamples).