Modeling and Estimation of Temporal Episode Patterns in Paroxysmal Atrial Fibrillation

Mikael Henriksson, Alba Martín-Yebra, Monika Butkuvienė, Jakob Gulddahl Rasmussen, Vaidotas Marozas, Andrius Petrėnas, Aleksei Savelev, Pyotr G. Platonov, and Leif Sörnmo, *Fellow, IEEE*

Abstract*—*Objective: **The present study proposes a model-based, statistical approach to characterizing episode patterns in paroxysmal atrial fibrillation (AF). Thanks to the rapid advancement of noninvasive monitoring technology, the proposed approach should become increasingly relevant in clinical practice.** ***Methods:*** **History-dependent point process modeling is employed to characterize AF episode patterns, using a novel alternating, bivariate Hawkes self-exciting model. In addition, a modified version of a recently proposed statistical model to simulate AF progression throughout a lifetime is considered, involving non-Markovian rhythm switching and survival functions. For each model, the maximum likelihood estimator is derived and used to find the model parameters from observed data.** ***Results:*** **Using three databases with a total of 59 long-term ECG recordings, the goodness-of-fit analysis demonstrates that the proposed alternating, bivariate Hawkes model fits SR-to-AF transitions in 40 recordings and AF-to-SR transitions in 51; the corresponding numbers for the AF model with non-Markovian rhythm switching are 40 and 11, respectively. Moreover, the results indicate that the model parameters related to AF episode clustering, i.e., aggregation of temporal AF episodes, provide information complementary to the well-known clinical parameter AF burden.** ***Conclusion:*** **Point process modeling provides a detailed characterization of the occurrence pattern of AF episodes that may improve the understanding of arrhythmia progression.**

***Index Terms*—Atrial fibrillation, point process modeling, alternating bivariate Hawkes model, maximum likelihood estimation, episode clustering.**

Manuscript received February 5, 2020; revised April 30, 2020; accepted May 14, 2020. Date of publication May 20, 2020; date of current version December 21, 2020. This work was supported in part by the Swedish Research Council (2016-03382), in part by the Research Council of Lithuania (S-MIP-17/81), in part by JGR, in part by the Danish Council for Independent Research (DFF7014-00074), and in part by the Villum Foundation (#8721). *(Corresponding author: Alba Martín-Yebra.)*

Mikael Henriksson and Leif Sörnmo are with the Department of Biomedical Engineering and Center for Integrative Electrocardiology, Lund University.

Alba Martín-Yebra is with the Department of Biomedical Engineering and Center for Integrative Electrocardiology, Lund University, Lund 22100, Sweden (e-mail: alba.martin@bme.lth.se).

Monika Butkuvienė and Andrius Petrėnas are with the Biomedical Engineering Institute, Kaunas University of Technology, Kaunas, Lithuania.

Jakob Gulddahl Rasmussen is with the Department of Mathematical Sciences, Aalborg University.

Vaidotas Marozas is with the Biomedical Engineering Institute and Department of Electronics Engineering, Kaunas University of Technology.

Aleksei Savelev is with the St. Petersburg State University.

Pyotr G. Platonov is with the Center for Integrative Electrocardiology, Lund University and Arrhythmia Clinic, Skåne University Hospital.

Digital Object Identifier 10.1109/TBME.2020.2995563

I. Introduction

ATRIAL fibrillation (AF) is a progressive disease often initially manifested by intermittent episodes terminating spontaneously. In paroxysmal AF, episode duration varies substantially, lasting from less than 30 s to 7 days. Considering that AF is a heterogenous disease associated with significant comorbidities in certain patients, but no demonstrable disease in others [1], [2], it is not surprising that progression from paroxysmal AF to sustained forms of AF, i.e., persistent, long-standing, and permanent, does not occur in all patients [3]. While little is known today about the role of temporal episode patterns in AF progression [4], long-term continuous monitoring is, most likely, the tool to provide such knowledge.

The problem of how to characterize AF episode patterns received certain attention around the turn of the millennium. At the time, emphasis was put on univariate statistical analysis of either the intervals between consecutive AF episodes ("interepisode intervals") [5]–[7], or the intervals between the onsets of consecutive AF episodes ("interdetection intervals") [8]. While it was speculated that information on episode patterns can be useful to predict outcome [9], e.g., by relating the degree of episode clustering to antiarrhythmic therapy [6], the clinical significance was never investigated.

Most of the "millennial studies" were based on the series of RR intervals produced by the AF detector in an implantable device. This approach offered continuous operation for a year or more with the potential to characterize AF progression, though constrained by a limited storage capacity. Using a small data set, initial results suggested that interepisode intervals could be described by a homogenous Poisson model, implying that interepisode intervals follow an exponential probability density function (PDF) [5]. This PDF model was, however, later discarded in favor of the Weibull PDF [6], [7] or the power law PDF [8] as the latter two PDFs were found more adequate for modeling of interepisode intervals.

"AF density" is one of the very few parameters proposed for characterizing the aggregation of AF burden in patients subject to year-long monitoring using an implantable device [10]; AF burden is the percentage of time spent in AF during the

monitored period. AF density is defined on the interval [0,1], where a value close to 0 indicates that AF burden is evenly distributed on a day-to-day basis throughout the monitored period and a value close to 1 that AF burden is confined to an interval much shorter than the monitored period. Recently, AF density was modified to account for episode duration to become suitable for characterization of day-long recordings on an RR interval basis [11]: While the modified parameter could distinguish different episode patterns, it was also found to be strongly correlated with AF burden ($r = -0.94$).

The above-mentioned PDF-based approach to characterizing AF interepisode or interdetection intervals rests on the assumption that episodes are statistically independent—an assumption that may be questioned since AF episodes tend to cluster [6]. Recently, a model-based approach to simulating AF progression during a lifetime was proposed, where AF episodes are assumed to be history-dependent [12]. The main components of that model are the rates of activation and recovery, defined by survival functions, resulting in two series of points in time together defining the onset and end of successive AF episodes in a simulated patient. Following AF initiation, episodes are brief to then become increasingly longer until AF progresses to a sustained form. By exploring various combinations of model parameter values, the authors expressed great expectations that simulations could provide better understanding of the key mechanisms determining long-term progression of AF. In that study, parameter estimation from observed data was not addressed, one reason being that patient monitoring over a lifetime is impractical.

The purpose of the present study is to explore a model-based approach to characterizing AF episode patterns. A variation of the bivariate Hawkes self-exciting point process model [13], [14] is proposed for history-dependent modeling of the alternating transition times from non-AF to AF and vice versa. In the model, a transition increases the likelihood of observing additional transitions in the near future, thus allowing clustered transition patterns to be modeled. The conditional intensity function, i.e., the function specifying the mean number of transitions in an interval conditional on the past, is defined by a relatively small number of parameters and therefore suitable for statistical inference. While the alternating, bivariate Hawkes model is novel, point process modeling as such has been considered in biomedical applications for modeling of neural spike activity [15], heartbeat dynamics [16], [17], and physical activity [18]. In addition to proposing a variation of the Hawkes model, the present study provides a point process interpretation of the AF progression model in [12] and proposes a minor modification that makes the model well-suited for analyzing data observed during days and weeks. In addition, the modification paves the way for statistical inference, solved using the maximum likelihood (ML) method.

The paper is organized as follows. Section II introduces the alternating, bivariate Hawkes model together with the related ML estimator. Section III describes the AF progression model and the related ML estimator. Section IV describes the approach taken to goodness-of-fit analysis for model validation and three different long-term ECG databases. Section VI presents the results from the goodness-of-fit analysis as well as results on the relationship between certain model parameters, the modified AF density, and AF burden.

II. Point Process Modeling

The temporal episode pattern in paroxysmal AF is modeled by two point processes: one accounting for transitions from sinus rhythm (SR) to AF occurring at times ("points") $t_{1,1}, t_{1,2}, \ldots$, and another accounting for transitions from AF to SR occurring at times $t_{2,1}, t_{2,2}, \ldots$; the first index indicates the type of transition and the second the transition number. For simplicity, SR and AF are assumed to alternate, although, in practice, a non-AF rhythm may very well replace SR. The onset of the first AF episode and the end of the last AF episode are assumed to be entirely contained in the observation interval $[0, T]$. Thus, the first transition is from SR-to-AF and the last from AF-to-SR.

A bivariate point process is associated with two counting processes $\{N_1(t), N_2(t), t \geq 0\}$ describing the number of transitions that have occurred up to but not including t,

$$N_m(t) = \sum_{k=1}^{\infty} \mathbb{1}_{[t_{m,k}<t]}, \quad m = 1, 2, \tag{1}$$

where $N_m(0) = 0$, $\mathbb{1}_{[\cdot]}$ is the indicator function which equals either 1 or 0 depending on whether the condition inside the square brackets is true or not, and k the index of past points. A bivariate point process is completely characterized by the two conditional intensity functions $\lambda_1(t)$ and $\lambda_2(t)$, defined by [19]

$$\lambda_m(t) = \lim_{\Delta t \to 0} \frac{\Pr(N_m(t + \Delta t) - N_m(t) = 1 | \mathcal{H}_t)}{\Delta t}, \tag{2}$$

where the numerator is the conditional probability of a transition occurring in the interval $[t, t + \Delta t]$, and $\mathcal{H}_t$ is the history of the bivariate point process, i.e., the transition times $t_{1,1}, t_{2,1}, t_{1,2}, \ldots$ that have occurred up to but not including t.

The conditional intensity function $\lambda_m(t)$ involves a set of parameters that can be estimated using the ML method. The parametric dependence is made explicit by collecting all the model parameters in a vector $\boldsymbol{\theta}$ and expanding the notation to become $\lambda_m(t; \boldsymbol{\theta})$. For a bivariate process, the likelihood function is given by [19]

$$\mathcal{L}(\boldsymbol{\theta}; \mathbf{t}) = \left[\prod_{m=1}^{2} \prod_{k=1}^{N_m(T)} \lambda_m(t_{m,k}; \boldsymbol{\theta}) \right] \cdot \exp\left(- \sum_{m=1}^{2} \int_0^T \lambda_m(t; \boldsymbol{\theta}) dt \right), \tag{3}$$

where the vector $\mathbf{t}$ contains the observed transition times.

The ML estimator is given by [20]

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} (\ln \mathcal{L}(\boldsymbol{\theta}; \mathbf{t})), \tag{4}$$

where the log-likelihood function is given by

$$\ln \mathcal{L}(\boldsymbol{\theta}; \mathbf{t}) = \sum_{m=1}^{2} \sum_{k=1}^{N_m(T)} \ln \lambda_m(t_{m,k}; \boldsymbol{\theta}) - \sum_{m=1}^{2} \int_0^T \lambda_m(t; \boldsymbol{\theta}) dt. \tag{5}$$

It should be noted that the approach taken to point process modeling does not explicitly account for AF episode duration, thus differing from the approaches in [6], [7].

A. Bivariate Hawkes Model

In the present study, the bivariate Hawkes model with exponential decays serves as the starting point for modeling episode patterns. The counting processes $N_1(t)$ and $N_2(t)$ have conditional intensity functions of the form [13]:

$$\lambda_m(t) = \mu_m + \sum_{n=1}^{2} \sum_{\{k:t>t_{n,k}\}} \alpha_{m,n} e^{-\beta_{m,n}(t-t_{n,k})}, \quad (6)$$

where $\mu_m > 0, \alpha_{m,n} \geq 0, \beta_{m,n} \geq 0$ for $m, n = 1, 2$. The main characteristic of the Hawkes model is that $\lambda_1(t)$ increases by $\alpha_{1,1}$ immediately after a transition ("self-excitation") and then decreases exponentially, defined by the decay parameter $\beta_{1,1}$, to the base intensity μ_1; the same characteristic applies to $\lambda_2(t)$ but then defined by $\alpha_{2,2}, \beta_{2,2}$, and μ_2. Thus, the process can exhibit a clustering behavior due to that the probability of additional transitions increases immediately after a transition. In addition to self-excitation, $\lambda_1(t)$ contains another term, defined by $\alpha_{1,2}$ and $\beta_{1,2}$, which lets $N_2(t)$ influence $N_1(t)$ ("cross-excitation"); $\lambda_2(t)$ is defined in the same way as $\lambda_1(t)$, but then by $\alpha_{2,1}$ and $\beta_{2,1}$. In the following, the model parameters are compactly represented by the vector $\boldsymbol{\theta}$,

$$\boldsymbol{\theta} = [\mu_1, \mu_2, \alpha_{1,1}, \beta_{1,1}, \alpha_{1,2}, \beta_{1,2}, \alpha_{2,1}, \beta_{2,1}, \alpha_{2,2}, \beta_{2,2}]. \quad (7)$$

B. Alternating, Bivariate Hawkes Model

Unfortunately, the bivariate Hawkes model does not impose alternating transitions, i.e., a transition from SR to AF is not necessarily followed by a transition from AF to SR to ensure that $t_{1,1} < t_{2,1} < t_{1,2} < t_{2,2} < \cdots$. Although the model parameters can be estimated and given an interpretation in physiological terms [18], the bivariate Hawkes model is not meaningful to use for simulating episode patterns as illustrated by Fig. 1(a). However, this disadvantage can be eliminated by multiplying each of the conditional intensity functions in (6) with a binary "occurrence" function, ensuring that AF occurs after SR,

$$o_1(t) = \begin{cases} 1, & N_1(t) = N_2(t - d_2), \\ 0, & \text{otherwise}, \end{cases} \quad (8)$$

and SR occurs after AF,

$$o_2(t) = \begin{cases} 1, & N_2(t) \neq N_1(t - d_1), \\ 0, & \text{otherwise}; \end{cases} \quad (9)$$

see Fig. 2 for an illustration of $o_1(t)$ and $o_2(t)$. The parameters d_1 and d_2 define the minimum duration of AF and SR, respectively. It should be noted that the choice of d_1 and d_2 has implications on the number of episodes contained in the data set: small values of d_1 and d_2 typically imply more episodes than do large values.

The conditional intensity functions producing an alternating, bivariate point process are given by

$$\tilde{\lambda}_m(t) = \lambda_m(t) o_m(t), \quad m = 1, 2. \quad (10)$$

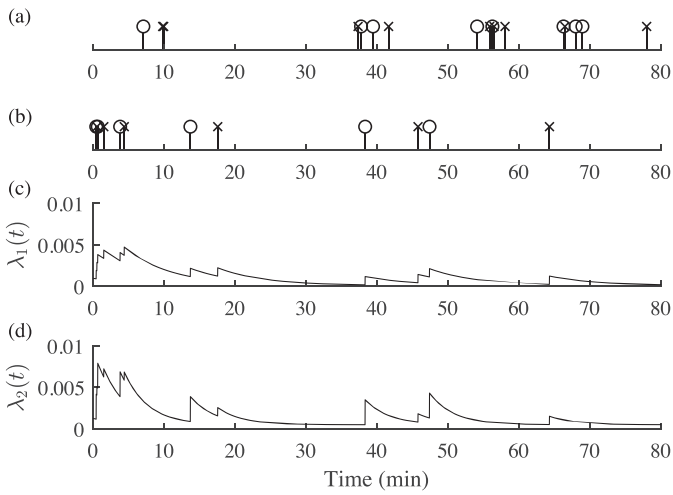

Fig. 1. (a) Simulated realization of the bivariate Hawkes point process, defined in (6), using $\boldsymbol{\theta} = [0.1, 0.5, 1, 2.6, 1, 2.6, 3, 5, 1, 5] \cdot 10^{-3}$. (b) Realization of the alternating, bivariate Hawkes point process, using $\boldsymbol{\theta}$ in (a), and related intensity functions (c) $\lambda_1(t)$ and (d) $\lambda_2(t)$. The marks "o" and "x" indicate SR-to-AF and AF-to-SR transitions, respectively. In (b)–(d), the minimum durations d_1 and d_2 are set to 0.

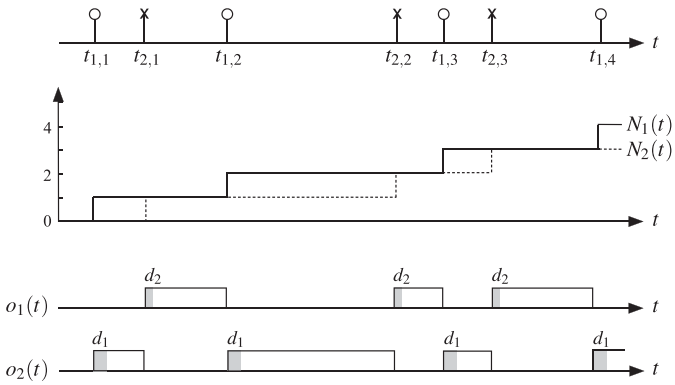

Fig. 2. Illustration of the alternating, bivariate Hawkes model. The transition times ("o," SR-to-AF; "×," AF-to-SR), the counting processes $N_1(t)$ and $N_2(t)$ (solid and dashed line, respectively), and the occurrence functions $o_1(t)$ and $o_2(t)$ are displayed from top to bottom. The minimum duration of AF and SR (shaded) are denoted d_1 and d_2, respectively.

Thus, the structure of $\tilde{\lambda}_m(t)$ is identical to the bivariate Hawkes process in (6), except that a transition from SR to AF must, once a certain time d_1 has elapsed, be followed by a transition from AF to SR, and so on. Fig. 1(b)–(d) illustrate the alternating, bivariate Hawkes model and related conditional intensity functions. Clustered and non-clustered episode patterns are illustrated in Fig. 3, where the clustered pattern in Fig. 3(a) is produced using smaller values of the β-parameters than in Fig. 3(b), i.e., $\beta_{1,1} = \beta_{1,2} = 2.5 \cdot 10^{-3}, \beta_{2,1} = \beta_{2,2} = 5 \cdot 10^{-3}$ vs. $\beta_{1,1} = \beta_{1,2} = \beta_{2,1} = \beta_{2,2} = 8 \cdot 10^{-3}$.

The log-likelihood function of the bivariate Hawkes model has been derived in [21], however, owing to the introduction of the occurrence functions $o_1(t)$ and $o_2(t)$, the log-likelihood function of the alternating, bivariate Hawkes model has a structure that differs substantially from the one in [21], given by (see

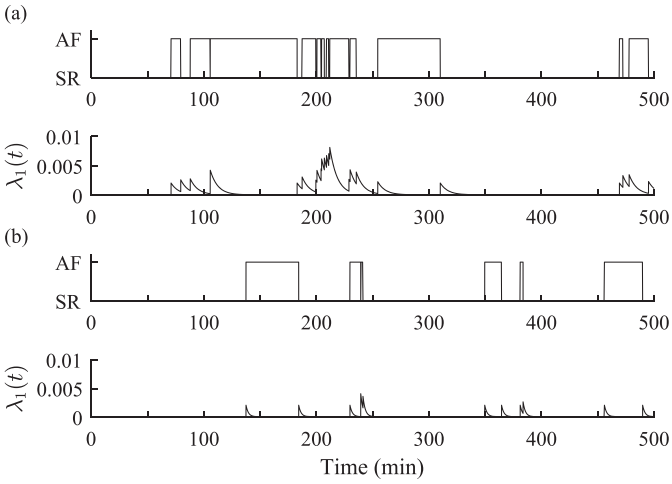

Fig. 3. Simulated (a) clustered and (b) non-clustered episode patterns using the alternating, bivariate Hawkes model using $\boldsymbol{\theta} = [0.1, 0.5, 2, 2.5, 2, 2.5, 1, 5, 2, 5] \cdot 10^{-3}$ and $[0.1, 0.5, 2, 8, 2, 8, 1, 8, 2, 8] \cdot 10^{-3}$, respectively. For reasons of clarity, $\lambda_1(t)$ is displayed rather than $\tilde{\lambda}_1(t)$. The minimum durations d_1 and d_2 are set to 0.

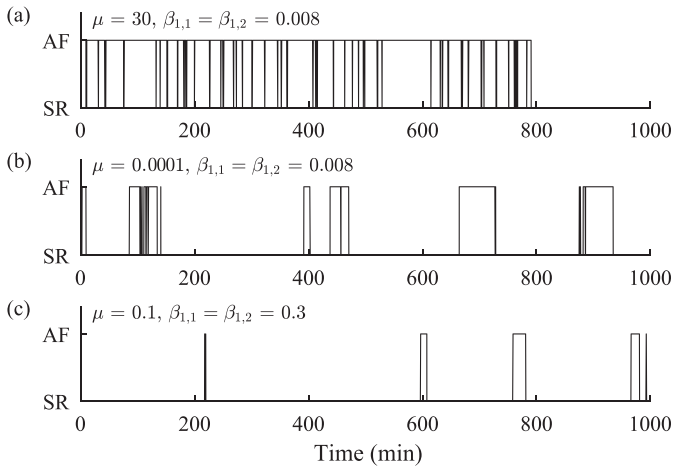

Fig. 4. (a) Episode pattern dominated by AF, simulated using the alternating, bivariate Hawkes model using $\boldsymbol{\theta} = [0.03, 0.001, 1 \cdot 10^{-6}, 0.008, 0.008, 0.008, 0.04, 0.07, 1 \cdot 10^{-6}, 0.07]$. (b) Episode pattern dominated by SR, generated using $\boldsymbol{\theta}$ in (a) but with μ_1 decreased to 0.0001. (c) Episode pattern dominated by SR, generated using $\boldsymbol{\theta}$ in (b) but with $\beta_{1,1}$ and $\beta_{1,2}$ increased from 0.008 to 0.3, leading to less clustering. The minimum durations d_1 and d_2 are set to 0.

Appendix A)

$$\ln \mathcal{L}_m^a(\boldsymbol{\theta}; \mathbf{t}) = - \sum_{\{k: t_{2,k+2-m} < T\}} \Bigg(\mu_m (\tau_{\tilde{m},k} - d_{\tilde{m}}) + \sum_{n=1}^{2} \frac{C_{m,n}(k)}{\beta_{m,n}} \left(e^{-\beta_{m,n} d_{\tilde{m}}} - e^{-\beta_{m,n} \tau_{\tilde{m},k}} \right) - \ln \left(\mu_m + \sum_{n=1}^{2} C_{m,n}(k) e^{-\beta_{m,n} \tau_{\tilde{m},k}} \right) \Bigg), \tag{11}$$

where $\tilde{m} = 2$ for $m = 1$ and, conversely, $\tilde{m} = 1$ for $m = 2$. Furthermore, $\tau_{1,k} = t_{2,k} - t_{1,k}$ is the duration of the k:th AF episode, and $\tau_{2,k} = t_{1,k+1} - t_{2,k}$ is the duration of the k:th SR "episode". The function $C_{m,n}(k)$ is defined by

$$C_{m,n}(k) = \alpha_{m,n} \sum_{l=1}^{k} e^{-\beta_{m,n}(t_{\tilde{m},k} - t_{n,l})}, \tag{12}$$

except for $m = n = 2$ when

$$C_{2,2}(k) = \alpha_{2,2} \sum_{l=1}^{k-1} e^{-\beta_{2,2}(t_{1,k} - t_{2,l})}. \tag{13}$$

The maximization of the log-likelihood function in (11) is performed using particle swarm optimization.

The base intensities μ_1 and μ_2 reflect the mean rates of SR-to-AF and AF-to-SR transitions, respectively, in the absence of self- and cross-excitation. The base intensity ratio

$$\mu = \frac{\mu_1}{\mu_2} \tag{14}$$

provides information on the dominating rhythm: $\mu > 1$ indicates dominance of AF (Fig. 4(a)) and $\mu < 1$ dominance of SR (Fig. 4(b)). Holding fixed μ and the other model parameter values used to generate the pattern in Fig. 4(b), the effect of increasing $\beta_{1,1}$ and $\beta_{1,2}$, leading to less clustering, is illustrated in Fig. 4(c). Moreover, for a fixed AF burden, it may be noted that a few long AF episodes occurring closely in time are characterized by a $\mu > 1$, whereas many short, evenly distributed AF episodes are characterized by a $\mu < 1$. Since the transition times form a series of alternating episode onsets and ends, the cross-excitation parameters $\alpha_{1,2}$ and $\alpha_{2,1}$ are usually much larger than the self-excitation parameters $\alpha_{1,1}$ and $\alpha_{2,2}$.

III. The AF Progression Model [12]

In its original formulation, the AF progression model was defined by non-Markovian rhythm switching and two rates, activation and recovery, which together determine the transition times. Interestingly, as shown below, this model can be interpreted as a point process related to the alternating Hawkes process. The two rates, denoted $\lambda_1^p(t)$ and $\lambda_2^p(t)$, can be viewed as conditional intensity functions.

Only the parameters accounting for AF-induced remodeling are considered as they are relevant for modeling patterns in day- to week-long recordings. Hence, "lifetime" parameters accounting for genetic disposition and age- and disease-related remodeling are omitted. Rather than using parameter values found in the literature or based on empiricism as was done in [12], the parameters are here estimated from observed data.

A. Interpretation of the Progression Model as a Point Process

The conditional intensity function $\lambda_1^p(t)$ is defined by four positive-valued parameters: the base intensity μ_{min}, the maximum intensity μ_{max}, the exponential rate β_{AF} with which $\lambda_1^p(t)$ increases to μ_{max} in AF, and the exponential rate β_{SR} with which $\lambda_1^p(t)$ decreases to μ_{min} in SR. During the k:th episode of either

AF or SR, $\lambda_1^p(t)$ is formulated as

$$\lambda_1^p(t) = \begin{cases} \mu_{\max} + \alpha_{\mathrm{AF},k} e^{-\beta_{\mathrm{AF}}(t-t_{1,k})}, & t_{1,k} \le t < t_{2,k}, \\ \mu_{\min} + \alpha_{\mathrm{SR},k} e^{-\beta_{\mathrm{SR}}(t-t_{2,k})}, & t_{2,k} \le t < t_{1,k+1}, \end{cases} \tag{15}$$

where

$$\alpha_{\mathrm{AF},k} = \lambda_1^p(t_{1,k}) - \mu_{\max}, \tag{16}$$

$$\alpha_{\mathrm{SR},k} = \lambda_1^p(t_{2,k}) - \mu_{\min}, \tag{17}$$

for $k \ge 1$ which guarantees continuity of $\lambda_1^p(t)$. The intensity function $\lambda_1^p(t)$ is initialized with $\mu_{\min}$. It is noted that $\lambda_1^p(t)$ increases in AF since $\alpha_{\mathrm{AF},k}$ is negative, whereas it decreases in SR since $\alpha_{\mathrm{SR},k}$ is positive. As before, the conditional intensity function accounting for transitions is given by

$$\tilde{\lambda}_1^p(t) = \lambda_1^p(t) o_1(t). \tag{18}$$

In [12], $\lambda_2^p(t)$ was assumed to approach 0 in AF to mimic long-term progression from paroxysmal to persistent AF. Since this behavior has no relevance for modeling episode patterns, $\tilde{\lambda}_2^p(t)$ is described by a simple Poisson-like model,

$$\tilde{\lambda}_2^p(t) = \mu_2^p o_2(t). \tag{19}$$

A similar model was previously considered in [5], but then for the purpose of characterizing interepisode intervals.

Collecting the model parameters in the vector $\boldsymbol{\theta}^p$,

$$\boldsymbol{\theta}^p = [\mu_{\min}, \mu_{\max}, \beta_{\mathrm{AF}}, \beta_{\mathrm{SR}}, \mu_2^p], \tag{20}$$

the log-likelihood function can be written as (see Appendix B)

$$\ln \mathcal{L}_1(\boldsymbol{\theta}^p; \mathbf{t}) = -\sum_{\{k: t_{2,k+1} < T\}} \Big((\mu_{\min}(\tau_{2,k} - d_2) + \frac{\alpha_{\mathrm{SR},k}}{\beta_{\mathrm{SR}}} \left(e^{-\beta_{\mathrm{SR}} d_2} - e^{-\beta_{\mathrm{SR}} \tau_{2,k}} \right) - \ln \left(\mu_{\min} + \alpha_{\mathrm{SR},k} e^{-\beta_{\mathrm{SR}} \tau_{2,k}} \right) \Big). \tag{21}$$

The function $\alpha_{\mathrm{SR},k}$ can be computed recursively, using

$$\alpha_{\mathrm{SR},k} = (\mu_{\max} - \mu_{\min})(1 - e^{-\beta_{\mathrm{AF}} \tau_{1,k}}) + e^{-\beta_{\mathrm{AF}} \tau_{1,k} - \beta_{\mathrm{SR}} \tau_{2,k-1}} \alpha_{\mathrm{SR},k-1}, \tag{22}$$

initialized by

$$\alpha_{\mathrm{SR},0} = 0. \tag{23}$$

In contrast to $\tilde{\lambda}_1(t)$ and $\tilde{\lambda}_2(t)$, $\tilde{\lambda}_1^p(t)$ and $\tilde{\lambda}_2^p(t)$ are decoupled from each other, and, consequently, ML estimation of $\boldsymbol{\theta}^p$ is decoupled from that of μ_2^p. The estimator of μ_2^p is given by

$$\hat{\mu}_2^p = \frac{N_2(T)}{\int_0^T o_2(t) dt}. \tag{24}$$

Simulated clustered and non-clustered AF episode patterns are illustrated in Fig. 5, where clustered patterns are associated with a smaller value of β_{SR}.

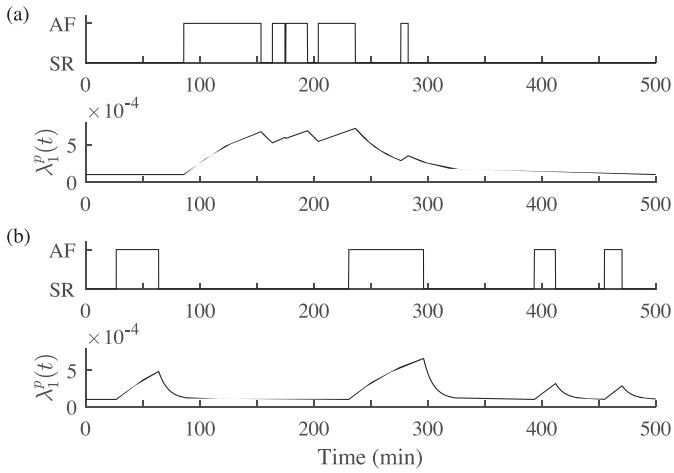

Fig. 5. (a) Clustered and (b) non-clustered episode patterns simulated using the AF progression model [12], using $\boldsymbol{\theta}^p = [0.1, 1, 0.25, 0.5, 0.5] \cdot 10^{-3}$ and $[0.1, 1, 0.25, 2, 0.5] \cdot 10^{-3}$, respectively. For reasons of clarity, $\lambda_1^p(t)$ is displayed rather than $\tilde{\lambda}_1^p(t)$. The minimum durations d_1 and d_2 are set to 0.

B. Relation Between Hawkes Model and AF Progression Model

This subsection highlights similarities and differences between the alternating, bivariate Hawkes model and the AF progression model, with emphasis put on $\lambda_1(t)$ and $\lambda_1^p(t)$ as these functions are history-dependent, while $\lambda_2^p(t)$ is not. To capture the full progression of $\lambda_1(t)$ and $\lambda_1^p(t)$, the occurrence function $o_1(t)$ is ignored.

The base intensities μ_1 and $\mu_{\min}$ define the lower bounds of $\lambda_1(t)$ and $\lambda_1^p(t)$, respectively, being approached in SR.

In the Hawkes model, the evolution of $\lambda_1(t)$ is completely determined by the two types of transitions, always decreasing to μ_1 regardless of whether the patient is in AF or SR. Immediately following a brief AF episode, a jump will occur in $\lambda_1(t)$ not only due to the AF-to-SR transition, defined by $\alpha_{1,2}$, but also the preceding SR-to-AF transition, defined by $\alpha_{1,1}$. Thus, a brief AF episode increases the likelihood of additional brief AF episodes. For small values of $\alpha_{1,1}$ or large values of $\beta_{1,1}$, there is no direct relationship between AF episode duration and $\lambda_1(t)$; instead, $\lambda_1(t)$ is related to the intervals between consecutive transitions to SR, i.e., similar to the "interdetection intervals" subject to analysis in [8].

On the other hand, in the AF progression model, no immediate jump in intensity occurs after a transition, but instead the intensity increases in AF and decreases in SR, with $\lambda_1^p(t)$ being a time-continuous function. Thus, contrary to $\lambda_1(t)$, $\lambda_1^p(t)$ will be larger following a longer AF episode. Unlike $\lambda_1(t)$ which is not upper bounded, $\lambda_1^p(t)$ is upper bounded by $\mu_{\max}$. If β_{AF}, describing the increase in $\lambda_1^p(t)$ during AF, is large enough, $\lambda_1^p(t)$ will equal $\mu_{\max}$ following every AF episode. This behavior resembles that of the Hawkes model, apart from that $\lambda_1(t)$ also depends on past transitions, contributing significantly when $\beta_{1,1}$ and $\beta_{1,2}$ are small enough.

In SR, both $\lambda_1(t)$ and $\lambda_1^p(t)$ decrease to their respective base intensity, given by $(\beta_{1,1}, \beta_{1,2})$ and β_{SR}, respectively. While the

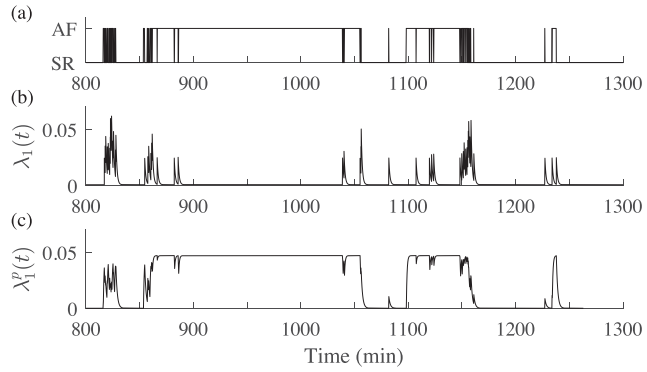

Fig. 6. (a) Clustered AF episodes in patient #28 of LTAFDB. The conditional intensity functions (b) $\lambda_1(t)$ and (c) $\lambda_1^p(t)$ are computed using the ML estimates of the respective model parameters.

overall behavior is similar in the two models, every past jump in intensity decreases separately in the Hawkes model, while β_{SR} describes the overall decay rate of the AF progression model. In both models, the decay parameters may be related to episode clustering since a slow decay increases the likelihood that an episode is followed by additional episodes.

Using patient data, Fig. 6 illustrates the main difference between the conditional intensity functions of the two models, namely that $\lambda_1(t)$ approaches μ_1 after a SR-to-AF transition, whereas $\lambda_1^p(t)$ approaches the upper bound μ_{max}.

IV. Goodness-of-Fit Analysis

The goodness-of-fit analysis explores how well the transition times of the AF episode pattern are modeled, accomplished by computing the integrated conditional intensity function for $\hat{\boldsymbol{\theta}}$,

$$\Lambda_{m,k} = \int_{t_{m,k-1}}^{t_{m,k}} \lambda_m(t;\hat{\boldsymbol{\theta}})dt, \quad m = 1, 2. \tag{25}$$

Invoking the time-scaling theorem, the integral in (25) transforms the pattern into a realization of a unit rate Poisson process [22]. To further facilitate model checking, another transformation, defined by $z_{m,k} = 1 - \exp(-\Lambda_{m,k})$, is used to produce independent, identically distributed uniform random variables in the interval (0,1]. The Kolmogorov–Smirnov (KS) plot is considered for model checking, displaying the cumulative distribution function (CDF) of $\Lambda_{m,k}$ versus the uniform CDF. The goodness-of-fit is judged by the distance from the diagonal line with which these two types of quantiles lie; a perfect model fit is obtained when all points appear on the diagonal. The fit is quantified by the largest distance ("KS distance") between the CDF of $\Lambda_{m,k}$ and the uniform CDF; this measure has previously been used to assess, e.g., the validity of point process models for heart rate variability analysis [16].

A point process model is deemed to fit the data when the maximum deviation between the two CDFs is within a 95% confidence interval.

V. Materials

The Physionet Long-Term AF Database (LTAFDB) consists of 84 24-h two-lead ambulatory ECG recordings acquired in patients with paroxysmal or persistent AF [23]. The beat-based annotation was automated, whereas the arrhythmia-based annotation resulted from manual review of the output of a commercial system for ECG analysis.

The MIT–BIH Atrial Fibrillation Database (AFDB) consists of 25 10-h, two-lead ambulatory ECG recordings from patients with AF, mostly paroxysmal [24]. The database was manually annotated with respect to type of beat, type of arrhythmia and episode transition times.

In addition, a database was acquired from patients with paroxysmal AF at the State University of St. Petersburg, Russia. The study was approved by the local ethical review board. The resulting database, named SPAFDB, consists of 37 three-lead ambulatory ECG recordings lasting from 1 to 7 days, amounting to a total of 160 days. Preliminary annotation of AF episodes in SPAFDB was performed using wavelet-based QRS detection [25], followed by AF detection based on fuzzy logic, involving information on ventricular rhythm, atrial rhythm, f wave morphology, and noise level [26]. Thereafter, manual review was performed to finalize the annotation process, with the aim to find undetected episodes, to discard falsely detected episodes, and, not the least, to establish the location of the episode transition times. The review was accomplished by an expert on AF analysis, consulting other experts in doubtful cases.

TABLE I
Model Parameters Subject to Estimation

Model	Parameter	Significance
Alternating, bivariate Hawkes model	μ_1	base intensity of $\lambda_1(t)$
	μ_2	base intensity of $\lambda_2(t)$
	$\alpha_{1,1}$	self-excitation of $\lambda_1(t)$
	β_1	intensity decay rate of $\lambda_1(t)$
	$\alpha_{1,2}$	cross-excitation of $N_2(t)$ on $N_1(t)$
	$\alpha_{2,1}$	cross-excitation of $N_1(t)$ on $N_2(t)$
	β_2	intensity decay rate of $\lambda_2(t)$
	$\alpha_{2,2}$	self-excitation of $\lambda_2(t)$
AF progression model	μ_{min}	base intensity of $\lambda_1^p(t)$
	μ_{max}	maximum intensity of $\lambda_1^p(t)$
	β_{AF}	exponential increase rate of $\lambda_1^p(t)$
	β_{SR}	exponential decrease rate of $\lambda_1^p(t)$
	μ_2^p	base intensity of $\lambda_2^p(t)$

VI. Results

The parameters d_1 and d_2 need to be set before ML estimation can be performed. Though d_1 may be set according to clinical guidelines (30 s) [27], the recent interest in brief, subclinical AF episodes [28], [29], providing an important reason why the present study was pursued, motivates the use of a small value of d_1 and, therefore, it is set to 3 s. Since d_2 has not received much clinical attention, it is set identical to d_1.

Assuming that $\beta_{1,1} = \beta_{1,2} = \beta_1$ and $\beta_{2,1} = \beta_{2,2} = \beta_2$, eight parameters have to be estimated in the alternating, bivariate Hawkes model. Table I lists the parameters subject to estimation in each of the two models.Maximum likelihood estimation is performed in ECG recordings with at least 10 episodes, i.e., 20 transitions. The transitions times are given by the database annotations. Table II provides a brief description of the analyzed databases.

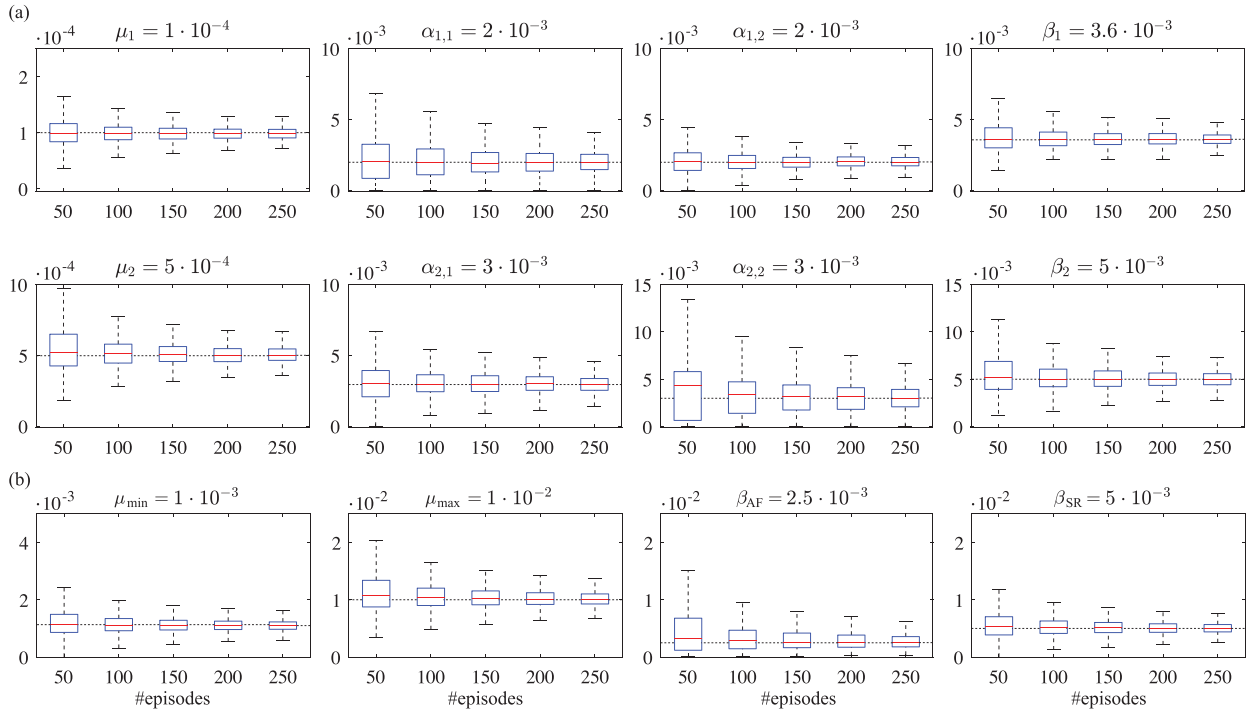

Fig. 7. Accuracy of model parameter estimates as a function of the number of observed episodes for (a) the alternating, bivariate Hawkes model and (b) the AF progression model. The box indicates the mean and the 25/75th percentiles, and the whiskers indicate ±2.7 times the standard deviation. The dotted horizontal line indicates the true parameter value (given in the diagram title).

TABLE II
DESCRIPTION OF ANALYZED DATABASES. THE MEDIAN #AF EPISODES PER RECORDING IS DENOTED E_m AND THE MEDIAN, MINIMUM, AND MAXIMUM EPISODE DURATION D_m, D_0, AND D_1, RESPECTIVELY. THE #RECORDINGS WITH AT LEAST 10 AF EPISODES IS DENOTED N_R

Database	N_R	#AF episodes	E_m	D_m, $[D_0, D_1]$ (s)
LTAFDB	30	1762	31	42, [3, 41759]
AFDB	6	210	32	100, [3, 8035]
SPAFDB	23	2293	72	31, [3, 117410]

TABLE III
THE KOLMOGOROV–SMIRNOV (KS) DISTANCE COMPUTED AS THE MEAN (STD) d_{KS} OF EACH OF THE THREE DATABASES. THE #RECORDINGS N_r FOR WHICH THE MODEL FITS THE DATA IS LISTED. N_r SHOULD BE RELATED TO N_R GIVEN IN TABLE II, I.E., 30/6/23 FOR LTAFDB/AFDB/SPAFDB, RESPECTIVELY

	LTAFDB		AFDB		SPAFDB	
Model	d_{KS}	N_r	d_{KS}	N_r	d_{KS}	N_r
$\tilde{\lambda}_1(t)$	0.23 (0.17)	19	0.23 (0.08)	4	0.14 (0.05)	17
$\tilde{\lambda}_1^p(t)$	0.25 (0.13)	18	0.24 (0.10)	4	0.15 (0.06)	18
$\tilde{\lambda}_2(t)$	0.17 (0.06)	24	0.15 (0.05)	5	0.11 (0.05)	22
$\tilde{\lambda}_2^p(t)$	$0.42^{\dagger}$(0.19)	8	0.30^{*}(0.11)	2	$0.49^{\dagger}$(0.22)	1

$^{\dagger}p < 0.001$ vs $\tilde{\lambda}_2(t)$; $^{*}p < 0.05$ vs $\tilde{\lambda}_2(t)$

A. Accuracy of ML Estimates

Using the models in Secs. II and III to simulate point processes, Fig. 7 presents the estimation performance as a function of the number of observed episodes for one set of parameter values. The results suggest that the ML estimates of both the alternating, bivariate Hawkes model and the AF progression model converge to the true values as the number of episodes increases, suggesting, in turn, estimator consistency. The results also suggest estimator unbiasedness since the true values lie approximately in the center of the boxes irrespective of the number of episodes. Several other sets of parameter values were also tested, yielding similar results.

B. Goodness-of-Fit Analysis Using LTAFDB/AFDB

Table III presents the results from the goodness-of-fit analysis. The alternating, bivariate Hawkes model $\{\tilde{\lambda}_1(t), \tilde{\lambda}_2(t)\}$ and the AF progression model $\tilde{\lambda}_1^p(t)$ are associated with about the same KS distance, denoted d_{KS}, on all databases, whereas $\tilde{\lambda}_2^p(t)$ is associated with a considerably larger distance. Using the Wilcoxon signed-rank test, d_{KS} of $\tilde{\lambda}_2^p(t)$ is significantly larger when compared to $\tilde{\lambda}_2(t)$.

Furthermore, Table III presents the number of recordings for which the model fits the data. The models $\tilde{\lambda}_1(t), \tilde{\lambda}_2(t)$, and $\tilde{\lambda}_1^p(t)$ fit the data in most recordings of the databases, while $\tilde{\lambda}_2^p(t)$ does not. In quantitative terms, $\tilde{\lambda}_1(t)$ and $\tilde{\lambda}_2(t)$ fit the data in 40 and 51 recordings, respectively, out of the total of 59 recordings, whereas the corresponding numbers for the AF progression model are 40 and 11.

To shed further light on the significance of the number of episodes in the estimation, LTAFDB and AFDB are merged and divided into two groups of recordings, those with <25 episodes and those with more. Fig. 8 shows that the mean d_{KS} was considerably lower for recordings with at least 25 episodes

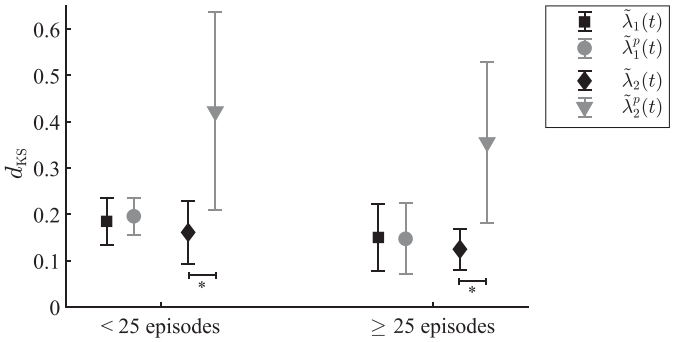

Fig. 8. Mean d_{KS} as a function of the number of AF episodes in different recordings. Intra-group differences were evaluated using Wilcoxon's signed-rank test (*$p < 0.005$).

than those with less than 25 episodes. It is noted that the mean d_{KS} was affected by the rather small number of episodes in many recordings (more than 50% of the recordings have less than 15 episodes).

C. Relationship Between AF Burden and Model Parameters

A desirable property of a model parameter is to convey information complementary to AF burden. As noted in Introduction, the modified AF density, denoted $\mathcal{A}$, was found to be strongly correlated with AF burden [11]. Here, using SPAFDB, a few model parameters of particular interest are analyzed with the aim to determine whether they are also strongly correlated with AF burden.

Fig. 9 presents scatter plots for AF burden and μ, the clustering parameters β_1 and β_{SR}, and $\mathcal{A}$. Both β_1 and β_{SR} are weakly correlated with AF burden ($r = 0.29$ and 0.27, respectively), whereas μ is somewhat more correlated ($r = 0.49$). The parameter $\mathcal{A}$ is strongly and negatively correlated to AF burden ($r = -0.73$) on SPAFDB, i.e., a result similar to that obtained using AFDB and LTAFDB ($r = -0.94$) [11].

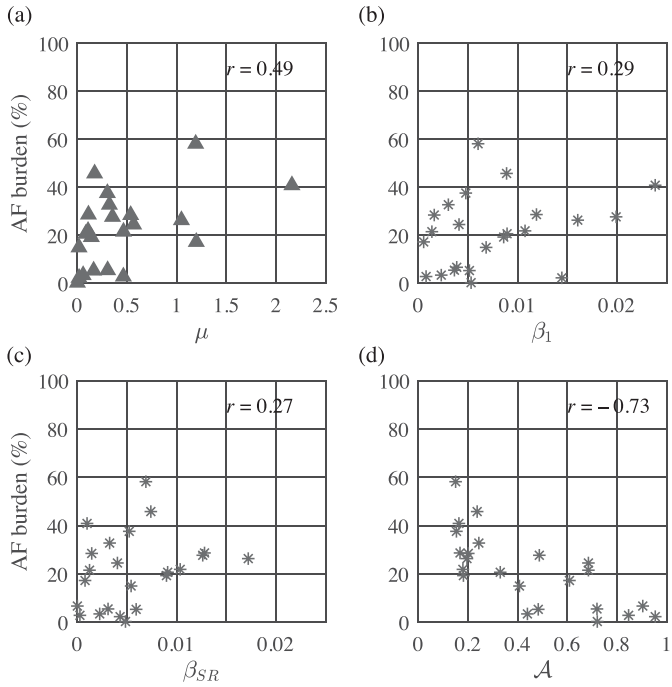

Fig. 9. Scatter plots of AF burden and (a) μ, (b) β_1, (c) β_{SR}, and (d) $\mathcal{A}$ [11]. The sample Pearson crosscorrelation coefficient r is given in each plot. Values deviating from the mean by more than five times the standard deviation are excluded.

VII. Discussion

Somewhat surprisingly, the recent clinical interest in long-term continuous monitoring [30], motivated by the need to assess risk of thrombus formation and ischemic stroke as well as to better understand disease progression, has not been paralleled by the development of methods which go beyond the analysis of AF burden. The present study attempts to address this lack by introducing a model-based, statistical approach to characterizing the dynamics of episode patterns, notably the degree of episode clustering, rate of transitions, and statistical coupling (cross-excitation) between SR-to-AF and AF-to-SR transitions. An advantage of this approach is the availability of methods for optimal parameter estimation, which for the considered models turned out to be analytically tractable, cf. the ML estimator in (11). Moreover, in contrast to a heuristic design approach, the statistical goodness-of-fit analysis provides information on how well the model fits the observed data. Yet another advantage is the possibility to simulate episode patterns—a feature which may, as pointed out in [12], serve as a testbed for experiments whose purpose is to investigate the relationship between the likelihood of episode initiation/termination and episode pattern appearance.

Studies investigating risk factors associated with AF have focused on whether AF is present or absent, sometimes together with information on AF burden [31]. The significance of AF burden as a predictor of patients at risk of ischemic stroke has been established in several studies [32]–[34]. A similar observation applies to minimal AF episode duration—a measure that may be related to increased risk of thrombus formation [34]. None of these measures say anything about whether episodes are clustered or distributed evenly during the monitored period, despite the fact that such information can be relevant for risk assessment of thrombus formation and ischemic stroke, as well as for better understanding arrhythmia progression. Since noninvasive, continuous monitoring has evolved to such a degree that a 14-day period or longer can be covered [35], the prerequisites for episode pattern analysis are becoming increasingly more favorable.

The results presented in Table III were obtained using $d_1 = 3$ s and thus subclinical AF episodes were included in the analysis. A comparison of the results in Table III to those obtained using a larger d_1 would be highly misleading as a larger d_1 means less episodes for analysis. Therefore, d_1 and d_2 should be set based on clinical considerations rather than on the results from goodness-of-fit analysis.

The insight made in the course of this work is that the AF progression model [12], originally developed for simulation

purposes only, can be interpreted as a point process, thus implying that a meaningful comparison of models can be made in terms of goodness-of-fit analysis. The results in Table III show that modeling of SR-to-AF transitions, defined by $\lambda_1(t)$ and $\lambda_1^p(t)$, are associated with approximately the same KS distance, suggesting that the different approaches to modeling episode clustering are quite similar. In contrast to $\lambda_2(t)$, modeling of AF-to-SR transitions by $\lambda_2^p(t)$ is considerably poorer. This result is in agreement with those presented in [6], [7] where the exponential distribution, originally proposed in [5], was found inadequate to model interepisode intervals.

The present approach requires the observation of a certain minimum number of AF episodes to obtain reasonably reliable parameter estimates. While the likelihood of observing more episodes increases as the length of the observation interval increases, the episode pattern may change its characteristics to such an extent that estimation is warranted in a segmented observation interval, using, e.g., a week-long segment length. In the present study, segmentation was not deemed necessary as the length of the recordings did not exceed one week. However, considering that month-long recordings can be acquired today using photoplethysmography-based smart wristbands [36]–[38], a segmentation strategy is probably necessary.

The model parameters are estimated from an observation interval defined by the onset of the first AF episode and the end of the last AF episode, not by the onset/end of the recording. This definition of observation interval may seem restrictive, however, the rationale is to avoid analyzing an episode in progress when the recording begins and/or an episode not terminated when the recording ends, i.e., episodes whose durations are unknown.

Contrary to the observation in [7] and later echoed in relation to the AF progression model [12], claiming that long episodes are immediately followed by a number of short episodes, visual inspection of the databases analyzed in the present study showed the reverse, namely that a long episode immediately preceded by a number of short episodes are much more frequent. This observation is particularly interesting since the alternating, bivariate Hawkes model better accounts for such a behavior as it is self-exciting (manifested by a larger than usual $\alpha_{1,1}$ and a smaller than usual β_1).

The alternating, bivariate Hawkes model can be used to improve the simulator recently proposed for generation of multilead ECGs in paroxysmal AF [39]. The simulator accounts for different characteristics such as switching between sinus rhythm and AF, repetition rate of f waves, varying P wave morphology, presence of atrial premature beats, and various types of noise. Instead of using the simple two-state, continuous-time Markov chain for rhythm switching, the proposed model can be used to simulate more realistic paroxysmal AF patterns. The three-state, continuous-time Markov chain used to model sinus rhythm, atrial flutter, and AF [40] may as well be replaced, but then probably by a trivariate Hawkes model.

The principal points of the present study are to introduce a model-based foundation to episode pattern characterization and to perform goodness-of-fit analysis. This analysis shows that $\lambda_1(t), \lambda_2(t)$ and $\lambda_1^p(t)$ are appropriate, although none of the models pretend to replicate behavior of underlying physiology. The correlation analysis in Section VI-C shows that the clustering parameters β_1 and β_{SR} are only weakly correlated with AF burden and, therefore, may provide complementary information. However, correlation analysis cannot say anything about the clinical significance of a certain parameter, thus necessitating a study which investigates the significance of the model parameters for the purpose of, e.g., predicting risk of stroke; unfortunately, none of the three databases of this study lends itself to such a study. Indeed, clinical evaluation of methods for episode pattern characterization has turned out to be a major challenge as neither the millennial studies [6]–[8] nor the study introducing AF density [41] investigated the relationship between pattern characteristics and patient outcome, although all were published in clinical journals. Considering the rapid development of monitoring technology, there are good reasons to develop methods for pattern characterization, exploring various principles, which are readily useful in future studies on patient outcome.

One idea to explore in a future study would be to investigate the extent by which an episode pattern reflects the degree of atrial electrical and structural remodeling. This idea draws on the observation that the course of pathophysiological processes underlying AF is commonly perceived to involve the development of structural abnormalities through inflammation-mediated replacement of atrial myocytes with fibrotic tissue, thinning of atrial walls, and atrial enlargement. Atrial structural remodeling is associated with changes in the clinical characteristics of AF, often manifested as episodes of increasing duration which are less likely to resolve spontaneously, ultimately deteriorating to a sustained form of AF.

Any method requires a certain minimum number of episodes to produce reliable results. By setting this number to 10, i.e., 20 transitions, a trade-off was made between the risk of model overfitting and the wish to include as many recordings as possible. In a future clinical study, this choice may very well be the subject of further investigation. The small data set is another limitation of the present study, however, the effort required to annotate week-long recordings should not be overlooked since annotation has to be performed also in recordings with too few episodes.

VIII. Conclusion

A model-based, statistical framework is proposed to characterize the pattern of SR-to-AF and AF-to-SR transitions, using an alternating, bivariate Hawkes model. Using ML parameter estimation techniques, the goodness-of-fit analysis demonstrates that the proposed model fits the data in the vast majority of recordings, implying that a wide range of episode patterns can be modeled. The proposed model offers a better overall fit to the data than the AF progression model thanks to better modeling of AF-to-SR transitions. The clinical significance of the model parameters remains to be investigated.

APPENDIX

A. Derivation of ML Estimator for the Alternating, Bivariate Hawkes Model

The starting point for deriving the estimator is the cumulative distribution function (CDF) $P(t_{1,k+1})$, which for $\lambda_1(t)$ in (6) is given by

$$P(t_{1,k+1}) = 1 - \Pr(t_{1,k+1} > t) = 1 - \exp\left[-\int_{t_{2,k}+d_2}^{t} \lambda_1(t')dt'\right]. \quad (26)$$

The integration interval accounts for the minimum duration of an SR episode, i.e., d_2, which allows the use of $\lambda_1(t)$ instead of $\tilde{\lambda}_1(t)$. Insertion of $\lambda_1(t)$ into (26) yields

$$\begin{aligned} P(t_{1,k+1}) &= 1 - \exp\left[-\int_{t_{2,k}+d_2}^{t} \left(\mu_1 + C_{1,1}(k)e^{-\beta_{1,1}(t'-t_{2,k})} + C_{1,2}(k)e^{-\beta_{1,2}(t'-t_{2,k})}\right)dt'\right] \\ &= 1 - \exp\left[-\mu_1(\tau_{2,k} - d_2) - \frac{C_{1,1}(k)}{\beta_{1,1}}\left(e^{-\beta_{1,1}d_2} - e^{-\beta_{1,1}\tau_{2,k}}\right) - \frac{C_{1,2}(k)}{\beta_{1,2}}\left(e^{-\beta_{1,2}d_2} - e^{-\beta_{1,2}\tau_{2,k}}\right)\right], \end{aligned} \quad (27)$$

where the functions $C_{1,1}(k)$ and $C_{1,2}(k)$ are defined in (12). The corresponding PDF $p(t_{1,k+1})$ is obtained by differentiating $P(t_{1,k+1})$ with respect to $t_{1,k+1}$, yielding

$$\begin{aligned} p(t_{1,k+1}) &= \left(\mu_1 + C_{1,1}(k)e^{-\beta_{1,1}\tau_{2,k}} + C_{1,2}(k)e^{-\beta_{1,2}\tau_{2,k}}\right) \\ &\quad \cdot \exp\left[-\mu_1(\tau_{2,k} - d_2) - \frac{C_{1,1}(k)}{\beta_{1,1}}\left(e^{-\beta_{1,1}d_2} - e^{-\beta_{1,1}\tau_{2,k}}\right) - \frac{C_{1,2}(k)}{\beta_{1,2}}\left(e^{-\beta_{1,2}d_2} - e^{-\beta_{1,2}\tau_{2,k}}\right)\right]. \end{aligned} \quad (28)$$

Applying the natural logarithm to $p(t_{1,k+1})$ and accounting for all $\tau_{2,k}$ yield the log-likelihood function in (11).

The PDF of $t_{2,k}$ is obtained in a similar manner, with μ_1, $\beta_{1,1}, \beta_{1,2}, C_{1,1}(k)$, and $C_{1,2}(k)$ replaced by the corresponding parameters and functions that define $\lambda_2(t)$.

B. Derivation of ML Estimator for the AF Progression Model

For the AF progression model, the CDF $P^p(t_{1,k+1})$ is expressed in terms of $\lambda_1^p(t)$,

$$\begin{aligned} P^p(t_{1,k+1}) &= 1 - \Pr(t_{1,k+1} > t) \\ &= 1 - \exp\left[-\int_{t_{2,k}+d_2}^{t} \lambda_1^p(t')dt'\right] \\ &= 1 - \exp\left[-\int_{t_{2,k}+d_2}^{t} \left(\mu_{\min} + \alpha_{\text{SR},k}e^{-\beta_{\text{SR}}(t'-t_{2,k})}\right)dt'\right] \\ &= 1 - \exp\left[-\mu_{\min}(\tau_{2,k} - d_2) - \frac{\alpha_{\text{SR},k}}{\beta_{\text{SR}}}\left(e^{-\beta_{\text{SR}}d_2} - e^{-\beta_{\text{SR}}\tau_{2,k}}\right)\right]. \end{aligned} \quad (29)$$

The corresponding PDF $p^p(t_{1,k+1})$ is

$$p^p(t_{1,k+1}) = \left(\mu_{\min} + \alpha_{\text{SR},k}e^{-\beta_{\text{SR}}\tau_{2,k}}\right) \cdot \exp\left[-\mu_{\min}(\tau_{2,k} - d_2) - \frac{\alpha_{\text{SR},k}}{\beta_{\text{SR}}}\left(e^{-\beta_{\text{SR}}d_2} - e^{-\beta_{\text{SR}}\tau_{2,k}}\right)\right], \quad (30)$$

which after applying the natural logarithm and accounting for all values of $\tau_{2,k}$ yield the log-likelihood function in (21).

REFERENCES

[1] C. B. de Vos *et al.*, "Progression from paroxysmal to persistent atrial fibrillation clinical correlates and prognosis," *J. Am. College Cardiol.*, vol. 55, pp. 725–731, 2010.

[2] G. J. Padfield *et al.*, "Progression of paroxysmal to persistent atrial fibrillation: 10-year follow-up in the Canadian registry of atrial fibrillation," *Heart Rhythm*, vol. 14, pp. 801–807, 2017.

[3] A. Shukla and A. B. Curtis, "Avoiding permanent atrial fibrillation: Treatment approaches to prevent disease progression," *Vascular Health Risk Manage.*, vol. 10, pp. 1–12, 2014.

[4] L. Mont and E. Guasch, "Atrial fibrillation progression: How sick is the atrium?" *Heart Rhythm*, vol. 14, pp. 808–809, 2017.

[5] G. S. Greer *et al.*, "Random and nonrandom behavior of symptomatic paroxysmal atrial fibrillation," *Am. J. Cardiol.*, vol. 64, pp. 339–342, 1989.

[6] A. M. Gillis and M. S. Rose, "Temporal patterns of paroxysmal atrial fibrillation following DDDR pacemaker implantation," *Am. J. Cardiol.*, vol. 85, pp. 1445–1450, 2000.

[7] W. F. Kaemmerer, M. S. Rose, and R. Mehra, "Distribution of patients' paroxysmal atrial tachyarrhythmia episodes: Implications for detection of treatment efficacy," *J. Cardiovascular Electrophysiol.*, vol. 12, pp. 121–130, 2001.

[8] L. A. Shehadeh, L. S. Liebovitch, and M. A. Wood, "Temporal patterns of atrial arrhythmia recurrences in patients with implantable defibrillators: Implications for assessing antiarrhythmic therapies," *J. Cardiovascular Electrophysiol.*, vol. 12, pp. 303–309, 2002.

[9] A. M. Gillis, "Modeling temporal patterns of atrial tachyarrhythmias: A new surrogate outcome measure for clinical studies?" *J. Cardiovascular Electrophysiol.*, vol. 13, pp. 310–311, 2002.

[10] E. I. Charitos *et al.*, "Atrial fibrillation density: A novel measure of atrial fibrillation temporal aggregation for the characterization of atrial fibrillation recurrence pattern," *Appl. Cardiopulmonary Pathophysiol.*, vol. 17, pp. 3–10, 2013.

[11] M. Šimaitytė *et al.*, "Quantitative evaluation of temporal occurrence patterns of paroxysmal atrial fibrillation," in *Proc. Comput. Cardiol.*, 2018, vol. 45, pp. 1–4.

[12] E. T. Y. Chang *et al.*, "A stochastic individual-based model of the progression of atrial fibrillation in individuals and populations," *PLoS ONE*, vol. 11, 2016, Art. no. e0152349.

[13] A. G. Hawkes, "Spectra of some self-exciting and mutually exciting point processes," *Biometrika*, vol. 58, pp. 83–90, 1971.

[14] J. G. Rasmussen, "Lecture notes: Temporal point processes and the conditional intensity function," 2018, *arXiv: 1806.00221v1* [stat.ME].

[15] W. Truccolo *et al.*, "A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects," *J. Neurophysiol.*, vol. 93, pp. 1074–1089, 2005.

[16] Z. Chen, E. N. Brown, and R. Barbieri, "Characterizing nonlinear heartbeat dynamics within a point process framework," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1335–1347, Jun. 2010.

[17] G. Valenza *et al.*, "Mortality prediction in severe congestive heart failure patients with multifractal point-process modeling of heartbeat dynamics," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 10, pp. 2345–2354, Oct. 2018.

[18] A. Paraschiv-Ionescu, E. Buchser, and K. Aminian, "Unraveling dynamics of human physical activity patterns in chronic pain conditions," *Sci. Rep.*, vol. 3, 2019, doi: 10.1038/srep02019.
[19] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. 2nd ed. Berlin, Germany: Springer, 2003.
[20] C. G. Bowsher, "Modelling security market events in continuous time: Intensity based, multivariate point process models," *J. Econometrics*, vol. 141, pp. 876–912, 2007.
[21] Y. Chen, "Likelihood function for multivariate Hawkes processes," 2016. [Online]. Available: http://www.math.fsu.edu/∼ychen/research/HawkesLikelihood.pdf
[22] E. N. Brown *et al.*, "The time-rescaling theorem and its application to neural spike train data analysis," *Neural Comput.*, vol. 14, pp. 325–346, 2002.
[23] S. Petrutiu, A. V. Sahakian, and S. Swiryn, "Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans," *Europace*, vol. 9, pp. 466–470, 2007.
[24] G. B. Moody and R. G. Mark, "A new method for detecting atrial fibrillation using R-R intervals," in *Proc. Comput. Cardiol.*, 1983, vol. 10, pp. 227–230.
[25] J. P. Martínez *et al.*, "A wavelet-based ECG delineator: Evaluation on standard databases," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, Apr. 2004.
[26] A. Petrėnas *et al.*, "Detection of occult paroxysmal atrial fibrillation," *Med. Biol. Eng. Comput.*, vol. 53, pp. 287–297, 2015.
[27] P. Kirchhof *et al.*, "2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS," *Europace*, vol. 18, pp. 1609–1678, 2016.
[28] A. C. Flint *et al.*, "Detection of paroxysmal atrial fibrillation by 30-day event monitoring in cryptogenic ischemic stroke: The stroke and monitoring for PAF in real time (SMART) registry," *Stroke*, vol. 43, pp. 2788–2790, 2012.
[29] J. O. Cerasuolo, L. E. Cipriano, and L. A. Sposato, "The complexity of atrial fibrillation newly diagnosed after ischemic stroke and transient ischemic attack: Advances and uncertainties," *Current Opinion Neurol.*, vol. 30, pp. 28–37, 2017.
[30] R. Lee and S. Mittal, "Utility and limitations of long-term monitoring of atrial fibrillation using an implantable loop recorder," *Heart Rhythm*, vol. 15, pp. 287–295, 2018.
[31] L. Y. Chen *et al.*, "Atrial fibrillation burden: Moving beyond atrial fibrillation as a binary entity. A scientific statement from the American Heart Association," *Circulation*, vol. 137, pp. e623–e644, 2018.
[32] G. Boriani *et al.*, "AF burden is important—fact or fiction?" *Int. J. Clin. Practice*, vol. 68, pp. 444–452, 2014.
[33] A. S. Go *et al.*, "Association of burden of atrial fibrillation with risk of ischemic stroke in adults with paroxysmal atrial fibrillation: The KP-RHYTHM Study," *JAMA Cardiol.*, vol. 3, pp. 601–608, 2018.
[34] R. Mahajan *et al.*, "Subclinical device-detected atrial fibrillation and stroke risk: A systematic review and meta-analysis," *Eur. Heart J.*, vol. 39, pp. 1407–1415, 2018.
[35] M. Yenikomshian *et al.*, "Cardiac arrhythmia detection outcomes among patients monitored with the Zio patch system: A systematic literature review," *Current Med. Res. Opinion*, vol. 35, no. 10, pp. 1659–1670, 2019.
[36] A. Sološenko *et al.*, "Detection of atrial fibrillation using a wrist-worn device," *Physiol. Meas.*, vol. 40, no. 2, 2019, Art. no. 025003.
[37] M. V. Perez *et al.*, "Large-scale assessment of a smartwatch to identify atrial fibrillation," *New Eng. J. Med.*, vol. 381, pp. 1909–1917, 2019.
[38] T. Pereira *et al.*, "Photoplethysmography based atrial fibrillation detection: A review," *NPJ Digit Med.*, vol. 3, Jan. 2020, doi: 10.1038/s41746-019-0207-9.
[39] A. Petrėnas *et al.*, "Electrocardiogram modeling during paroxysmal atrial fibrillation: Application to the detection of brief episodes," *Physiol. Meas.*, vol. 38, pp. 2058–2080, 2017.
[40] M. Masè *et al.*, "Modeling framework for the generation of synthetic RR series during atrial arrhythmias," in *Proc. IEEE Conf. Eng. Med. Biol.*, 2019, vol. 41, pp. 6347–6350.
[41] E. I. Charitos *et al.*, "A comprehensive evaluation of rhythm monitoring strategies for the detection of atrial fibrillation recurrence: Insights from 647 continuously monitored patients and implications for monitoring after therapeutic interventions," *Circulation*, vol. 126, pp. 806–814, 2012.