Atrial Fib

Tuesday, 8 April 2025 11:28

Detection of Atrial Fibrillation (AF)

- Research on detecting AF episodes has spanned over 30 years.
- Despite extensive work, a fully robust AF detector for continuous long-term ECG recordings and handheld device recordings remains undeveloped.
- A major challenge is the **high false alarm rate**, caused by:
  - Ectopic beats,
  - Noisy signal segments,
  - Non-AF arrhythmias that mimic AF rhythm patterns (see [1]).
- Human detection of AF relies on three core features:
  1. **Highly irregular rhythm**
  2. **Absence of P waves**
  3. **Presence of f waves**
- These characteristics are often incorporated into the design of AF detection algorithms, but with varying effectiveness.

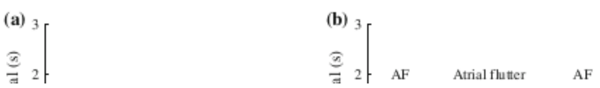
RR interval: tid mellan QRS komplex

Characterizing Rhythm Irregularity for AF Detection

- Translating **"highly irregular rhythm"** into a usable detection parameter is difficult due to the limited knowledge on effective features for characterizing irregularity.
- Numerous detection parameters have been proposed, each targeting a specific aspect of rhythm irregularity.
- **Early research** questioned whether RR intervals during AF were **random or deterministic** [2].
  - Results showed nonzero correlation between observed and predicted RR intervals, suggesting some predictability.
  - However, these findings were not consistent across all patients, making **correlation-based parameters unreliable** for AF detection.
- Another study found that RR intervals during AF show a **white noise-like spectrum** when analyzed minute-by-minute [3], suggesting randomness.
- **Heart rate** tends to be higher during AF episodes than during sinus rhythm.
  - Though **heart rate alone** is not a reliable detector, integrating it with rhythm irregularity parameters can **boost detection accuracy**.
  - Heart rate is usually represented by the **mean RR interval** within the detection window.
- **Confounding arrhythmias** complicate AF detection as they may present **RR interval patterns similar to AF**, increasing false alarms.
  - Important sources of false alarms include:
    - **Ventricular premature beats (VPBs)**
    - **Atrial premature beats (APBs)**
    - **Atrial flutter**
    - **Bigeminy**
    - **Trigeminy**
  - These are illustrated in **Fig. 4.1** showing representative RR interval series.
- **QRS detection inaccuracies** (due to muscle noise, motion artifacts, or large T waves) also contribute to false alarms.
- Shorter **detection windows**, necessary for detecting brief AF episodes, **increase the risk** of misclassifying non-AF rhythms.

Incorporating Waveform and Signal Quality in AF Detection

- When using information about **P waves** and/or **f waves** for AF detection, it is crucial to also include **signal quality** indicators.
  - This ensures that unreliable or noisy wave measurements do not **disrupt detection performance**.
  - Continuous long-term ECG recordings, often used in clinical studies, tend to have **variable noise levels**, making signal quality assessment essential for **trustworthy analysis**.
- The clinical definition of an AF episode is a **minimum duration of 30 seconds**:
  - This criterion was published in the **ACC/AHA/ESC 2006 guidelines** for AF patient management [4].
  - Although the reasoning behind the 30-second threshold was not thoroughly explained, the guidelines noted shorter episodes may still be relevant in cases involving:
    - **Symptomatic patients**
    - **Pre-excitation**
    - **Assessment of therapeutic interventions**
- In later updates:
  - The **2014 guidelines [5]** did **not** mention a minimum duration.
  - The **2016 guidelines [6] reinstated** the 30-second minimum, aligning again with the 2006 definition.



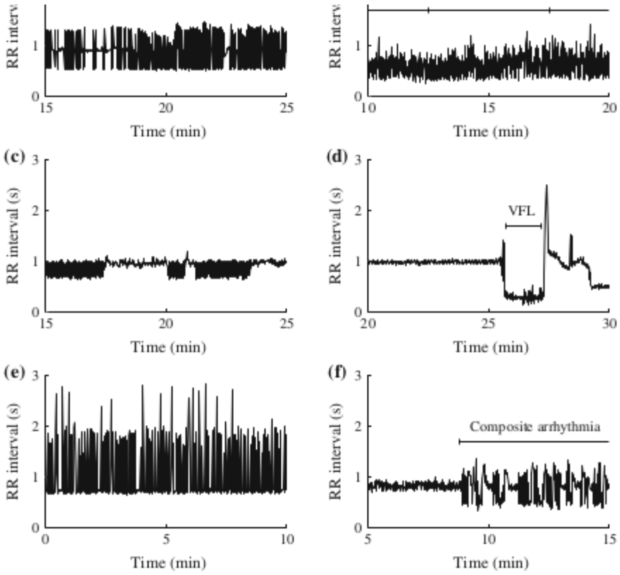


Fig. 4.1 – RR Interval Patterns Confounding AF Detection

- This figure presents RR interval patterns that can lead to **false AF detections**, illustrating the challenge of distinguishing AF from similar arrhythmias.
- All examples are from the **MIT–BIH Arrhythmia Database**.

(a)

- **Multiple ventricular premature beats**
  - Includes **bigeminy** and **trigeminy** patterns.
  - Irregular but structured rhythm may resemble AF.

(b)

- **Atrial flutter surrounded by AF episodes**
  - Clear alternation between **AF** and **atrial flutter**, making transitions difficult to classify.

(c)

- **Second degree atrioventricular (AV) block**
  - Regular interruptions in RR intervals due to dropped beats.

(d)

- **Ventricular flutter (VFL)** episode
  - Sudden shift into VFL with rapid, very short RR intervals.

(e)

- **Sinus bradycardia**
  - Slow, fairly regular rhythm with occasional outliers.

(f)

- **Composite arrhythmia episode**
  - Involves a mix of:
    - **AF**
    - **Atrial flutter**
    - **Atrial bigeminy**
    - **Supraventricular tachycardia**
    - **Atrioventricular junctional rhythm**
    - **Atrial premature beats**
  - Complex and irregular rhythm presents significant detection challenges.

These diverse patterns emphasize the need for **robust AF detectors** that can handle rhythm variability and differentiate AF from other irregular arrhythmias.

Clinical Significance of AF Episodes Shorter Than 30 Seconds

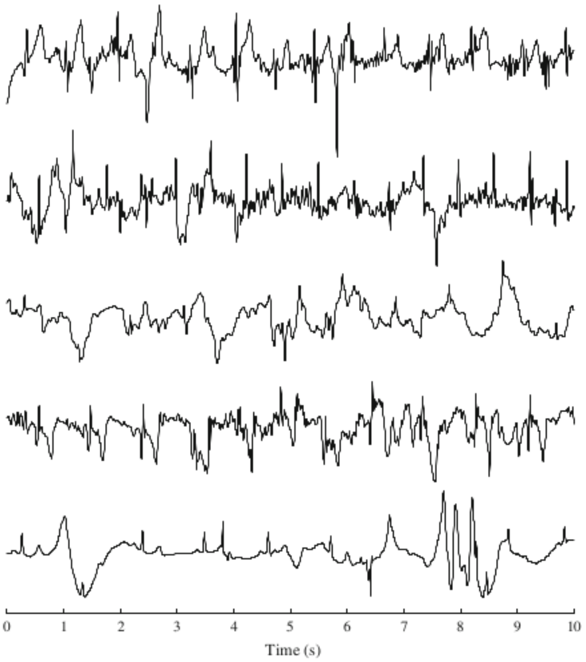
- Recent clinical research has **increased focus on AF episodes shorter than 30 seconds**, particularly regarding their **relation to stroke risk**.
- These short episodes might be **linked to atrial thrombus formation** and may act as **biomarkers** for longer, undetected episodes [12–14].
- **AF burden** is a key concept, defined as:

  > ***The proportion of the total recording time that a patient is in AF.***

  - A patient with **many short AF episodes** can have a **higher AF burden** than someone with fewer but longer episodes.
  - This may lead to a **higher thromboembolic risk**, even if no individual episode exceeds 30 seconds [15, 16].
- There is currently **no established minimum episode duration** that consistently conveys clinically significant information.

Challenges in Detecting Short AF Episodes:

- Long-term AF monitoring demands **automated event detection** for practical use.
- Detector design often imposes a **lower detection limit**, preventing recognition of very brief episodes:
  - For example, **RR interval histogram analysis** requires a large number of intervals, excluding episodes under two minutes.
  - Many **ECG-based detectors** cannot detect episodes <2 minutes.
  - Some **implantable devices** only detect episodes **≥6 minutes** [17, 18].
- Despite clinical studies showing that **episodes <30s are relevant**, most detectors:
  - Are **commercially developed** and proprietary.
  - Lack **published performance data** for short episodes [12, 13, 19, 20].
- Consequently, there is a pressing need to **design and validate new AF detectors** capable of:
  - Identifying **very brief episodes**,
  - Enabling further study of their **clinical impact**.



Variability in AF Episode Duration and Challenges in Mobile Detection

- **AF episode durations vary widely**, from **<30 seconds to over 7 days**.
  - Episodes longer than 7 days are termed **persistent AF** [4].
- AF detection should **avoid rigid assumptions** regarding:
  - **Episode duration**
  - **Minimum distance between AF episodes**
  - Similar to QRS detection strategies, a **least-informative approach** is recommended to preserve **clinically relevant data** [21].
- **Merging episodes** separated by just a few seconds could cause **important information loss**.

Impact of Handheld and Smartphone Devices:

- These devices **enable detection** of previously **unidentified AF** [22–29].
- However, they introduce **challenges due to lower signal quality**, often **inferior to clinical modalities**.
  - Refer to **Fig. 4.2** for examples of poor signal quality.
- Handheld/smartphone devices typically:
  - Use **a single lead**, which may **not reliably capture atrial activity**.
  - Depend primarily on **rhythm-based detection**.
  - Consider **f and P wave morphology** as **supplementary information**, rather than core detection criteria.

Overview of AF Detection Design Principles and Chapter Structure

- The chapter reviews **main design principles** in AF detection, focusing on:
  - **Rhythm-only analysis** using **RR interval series** (see Sect. **4.2**),
  - Combined use of **rhythm and atrial wave morphology** (see Sect. **4.3**).
- Additional topics covered:
  - **Detector implementation aspects** (Sect. **4.4**),
  - **Performance measures** used in AF detection (Sect. **4.5**).
- While performance is mentioned throughout the chapter, **Sect. 4.6** is dedicated to it as the **main theme**, discussing:
  - Key **considerations** when evaluating detector performance.

- The chapter concludes in **Sect. 4.7** with:
  - A discussion on **ECG-derived information** types that may enhance **detection performance**.

RR Interval-Based Detection and Design Limitations

- Due to **low SNR**, detecting **P and f waves** is difficult, especially in non-invasive recordings.
- As a result, most AF detectors rely solely on **RR interval irregularity**, characterized by:
  - **Randomness**
  - **Variability**
  - **Complexity**
- **Rhythm-based detectors** dominate because:
  - They consume **less power** than those using **morphological features**.
  - This is crucial for **implantable devices**, which often cannot use morphology-based information.

Detector Design Trends:

- Detector designs have historically followed **ad hoc principles**, using:
  - One or a few simple parameters fed into a **basic classifier**.
  - These often **lack statistical or physiological grounding**.
- Despite this, **performance (sensitivity/specificity)** has improved, as shown in **Table 4.1**.
- Continued **performance improvement** is needed to reduce **false alarms** from:
  - **Ectopic beats**
  - **Non-AF arrhythmias**
  - **Noisy signals**

Input Data and Processing Techniques:

- Besides the RR interval series $x(0), \ldots, x(N-1)$, detectors may also use the **first difference**:

$$\Delta x(n) = x(n) - x(n-1), \quad n = 1, \ldots, N-1 \tag{4.1}$$

- N: Number of RR intervals (not ECG samples)
- A **sliding time window** approach is commonly used:
  - Typical ECG recordings are short (10–20s),
  - Detection parameters are recalculated as the window slides forward in time.
  - **One-RR interval sliding** offers best onset/offset resolution,
  - But **larger "slides"** (e.g., 50 intervals) may be used to **reduce computation time** [36].

Performance of Rhythm-Based AF Detectors (Table 4.1 Overview)

- This section outlines **key rhythm-based AF detection methods**, evaluated using the **MIT–BIH Atrial Fibrillation Database (AFDB)** or its subset **AFDB₁**.
- **Performance metrics** include:
  - **Sensitivity (Se)** – the ability to correctly detect AF episodes,
  - **Specificity (Sp)** – the ability to correctly reject non-AF episodes.
- **Summary of methods and results:**

Method by	Year	Database	Se (%)	Sp (%)
Tateno and Glass [31]	2001	AFDB	94.4	97.2
Dash et al. [32]	2009	AFDB₁	94.4	95.1
Lian et al. [33]	2011	AFDB	95.8	96.4
Lake and Moorman [34]	2011	AFDB	91.0	94.0
Huang et al. [35]	2011	AFDB	96.1	98.1
Shouldice et al. [36]	2012	AFDB	92.0	96.0
Lee et al. [37]	2013	AFDB₁	98.2	97.7
Zhou et al. [38]	2014	AFDB	96.9	98.3
Asgari et al. [39]	2015	AFDB	97.0	97.1
Petrėnas et al. [40]	2015	AFDB	97.1	98.3
Zhou et al. [41]	2015	AFDB	97.4	98.4

- Notes:
  - **AFDB₁** excludes records 4936 and 5091 due to incorrect annotations.
  - All detectors listed use **fixed-window computations**, but these can be replaced with **sliding windows** for improved temporal resolution.
  - **Additional rhythm-based detectors** not listed were not evaluated on AFDB [42–47].

4.2.1 Irregularity Parameters

- **Table 4.2** lists parameters used in AF detector design, grouped into five categories:
  1. **Statistical dispersion**
  2. **Entropy**
  3. **Symbolic dynamics**
  4. **Poincaré plot-based parameters**
  5. **Time-varying coherence function**
- Among these, **statistical dispersion parameters** are the most commonly used:
  - Examples:
    - **Root mean square of successive differences**
    - **Mean of absolute successive differences**
    - **Coefficient of variation**
- Detector decision strategies:
  - Some use **a single parameter** with **simple thresholding**.
  - Others use a **combination of parameters** fed into a **classifier**.
- Certain parameters are tied to **specific statistical tests**, such as:

- **Number of turning points**, which may be described alongside the parameter instead of being placed in **Sect. 4.2.6**, where classifiers are generally discussed.

Statistical Dispersion Parameters in Rhythm-Based AF Detection

- One commonly used measure is the **Coefficient of Variation (CV)** of the RR interval series $x(n)$, defined as:

$$P_{\text{CV}} = \frac{\sigma_x}{m_x} \tag{4.2}$$

  where:

  - σ_x: standard deviation of $x(n)$
  - m_x: mean of $x(n)$
- **Interpretation**:
  - Reflects **dispersion** in RR intervals.
  - Also captures changes in **heart rate**, since AF episodes are associated with **RR interval shortening**, reducing m_x.
  - Using $\Delta x(n)$ instead of $x(n)$ makes $m_{\Delta x}$ small and unstable, so m_x is used to ensure stability.
- **Performance Note**:
  - Detectors using either $x(n)$ or $\Delta x(n)$ in P_{CV} showed **comparable performance** [31].
- Another key parameter is the **Root Mean Square of Successive Differences (RMSSD)**:

$$P_{\text{RMSSD}} = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N-1}\Delta x^2(n)} \tag{4.3}$$

Table 4.2 – Parameters Used in Rhythm-Based AF Detection

Detection Parameter	**Publication**
Coefficient of variation	[31, 48]
Root mean square of successive differences	[32, 37]
Normalized mean of absolute successive differences	[48]
Number of turning points	[32]
Histogram-based parameters	[31, 35]
Entropy	
Shannon entropy	[32, 37, 38]
Sample entropy	[34, 49]
Simplified sample entropy	[40]
Symbolic Dynamics	[38, 41]
Poincaré Plot-Based	
$x(n)$ vs $x(n-1)$ + bin count	[50]
$\Delta x(n)$ vs $\Delta x(n-1)$ + bin count	[51]
$x(n)$ vs $\Delta x(n-1)$ + bin count	[33]
Time-Varying Coherence Function	[37]

These parameters form the **core set** used to quantify RR interval irregularity and classify AF episodes.

Normalized Dispersion Measures for AF Detection

- The **RMSSD** parameter (from Eq. 4.3) does **not reflect heart rate changes**.
  - To address this, a **heart rate-dependent threshold** can be used [32].
  - Alternatively, RMSSD can be interpreted as **heart rate-normalized**, using $\Delta x(n)$ in place of $x(n)$.
  - This makes the test **equivalent to** P_{CV}, but with mean and standard deviation computed over $\Delta x(n)$.

- A related parameter is the **Normalized Mean of Absolute Successive Differences (NMASD)** [48], defined as:

$$P_{\text{NMASD}} = \frac{1}{N-1}\sum_{n=1}^{N-1}\frac{|\Delta x(n)|}{m_x} \tag{4.4}$$

  where:

  - m_x: mean of $x(n)$
  - $\Delta x(n) = x(n) - x(n-1)$

Key Insights:

- The use of P_{NMASD} instead of P_{CV}, when based on $\Delta x(n)$, is somewhat redundant:
  - P_{NMASD} is effectively an **approximation** of P_{CV}.
  - Their **detection performance is nearly identical** [48].
- Conclusion:
  - The three parameters P_{CV}, P_{RMSSD}, and P_{NMASD} (Eqs. 4.2–4.4) **convey similar information**.
  - All reflect **RR interval dispersion**.

- A different detection parameter, discussed next, will provide similar information within the **Poincaré plot** framework.

Number of Turning Points

- The **turning point test** is a **nonparametric statistical test** used to evaluate whether a time series behaves like a sequence of **independent and identically distributed (i.i.d.) random variables**.
  - In a fully random series, any 3-sample combination is equally likely to occur.
  - A **turning point** is when the **middle value** is a **local maximum or minimum**.
  - Probability of a turning point in a random 3-sample sequence is **2/3**.
- For a series with N samples:
  - The **expected number of turning points** m_{TP} and the **standard deviation** σ_{TP} are:

$$m_{\mathrm{TP}} = \frac{2(N-2)}{3} \tag{4.5}$$

$$\sigma_{\mathrm{TP}} = \sqrt{\frac{16N-29}{90}} \tag{4.6}$$

- If the **observed number of turning points** N_{TP} falls **outside** the 95% confidence range:

$$m_{\mathrm{TP}} \pm 1.96\sigma_{\mathrm{TP}},$$

  then the series is **not completely random**.

Use in AF Detection

- Turning point analysis is used to **characterize RR interval irregularity** [32].
- Instead of a strict statistical test, **optimized limits** are applied to enhance **sensitivity and specificity**.
- If turning points fall **outside optimized limits**, the RR intervals may be **periodic**, e.g., due to **respiratory-modulated sinus rhythm**.
- However, **limitations exist**:
  - RR intervals in AF are not fully random—they exhibit **correlations** [2].
  - The test may **lose power** in detecting randomness in these cases.
  - It can also cause **false alarms** in the presence of **ectopic beats** or rapid rhythm changes.
  - Therefore, it's **less suitable** as a standalone method for AF detection.

Histogram-Based Parameters

- **RR interval histograms** differ in shape between **sinus rhythm** and **AF**, making them useful for AF detection.
- For histograms to be **representative**, they require a **large number of RR intervals**:
  - **~100 beats minimum** is often cited [31, 35].
- This large requirement creates a **trade-off**:
  - **More data** improves histogram accuracy,
  - But it **reduces temporal resolution**, limiting the ability to detect **short AF episodes**.
- If **fewer or wider bins** are used to shorten the detection window:
  - **Discrimination decreases** between rhythm types,
  - Accuracy of **AF onset and end detection** suffers.

Detection Approaches:

1. **Heuristic Feature-Based:**
   - Uses simple descriptors like:
     - **Height of histogram**
     - **Number of non-empty bins**
   - In AF:
     - Histogram is **broader**, with **lower peak** and **more non-empty bins** than in sinus rhythm.
   - Problems arise if **heart rate changes** within the window:
     - Sinus rhythm histogram may **broaden**, resembling AF.
     - Mitigated by using a **ΔRR interval histogram**:
       - Removes slow trends,
       - Narrows histogram spread for better resolution.

2. **Template Comparison Approach:**
   - Detection window's RR histogram is compared with **template histograms**, each corresponding to a **mean RR interval length range** [31].
   - Example template intervals: 350–399 ms, 1100–1149 ms, in **50 ms steps**.
   - **Windows outside** the predefined ranges are **excluded**.
   - Templates are ideally built from **large databases**, ensuring histograms reflect the **true underlying probability distribution** (PDF).
   - The same procedure can also be applied to **ΔRR histograms**.

Key Limitation:

- The **need for long windows** in histogram-based detection inherently **limits its use** for identifying **brief AF episodes**.

Statistical Comparison of Histograms in AF Detection

- In AF detection, the observed RR interval histogram (from a **sliding window**) is compared to **template histograms** [31].
- The **Kolmogorov–Smirnov (K–S) test**, a nonparametric method, is used to assess whether:
  - The observed RR intervals and template intervals belong to the **same distribution** [55].

- The **K–S statistic** is:
  - The **maximum vertical distance** between the **cumulative histograms** of the observed and template data (see **Fig. 4.3**).
  - Suitable for detecting **global distribution differences**, but **less sensitive to local differences**, such as:
    - Changes in the **number of peaks**.
- If the difference between distributions is **localized** (e.g., bimodal vs unimodal), the **Anderson–Darling test** may be more appropriate than K–S, as it uses **squared deviations** across the entire range [55].

Performance Evaluation:

- **Using RR series with the K–S test** (from [31]):
  - **Sensitivity:** 66.3%
  - **Specificity:** 99.0%
- **Using ΔRR series instead**:
  - **Sensitivity improved** to **94.4%**
  - **Specificity slightly decreased** to **97.2%**
- Authors did not explain the improvement, but a possible reason:
  - The **ΔRR histogram** may be **more unimodal** than the RR histogram,
  - Making it more suitable for the **K–S test's global difference metric**.

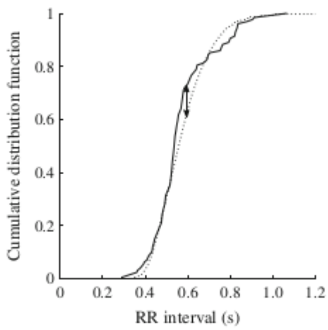


Fig. 4.3 Insight:

- Illustrates two cumulative histograms of RR intervals in AF.
- The **maximum distance** between them, which is used in the K–S test, is marked with an arrow.

Advanced Histogram-Based Approaches in AF Detection

- The **multi-template histogram approach** enhances characterization of RR interval distributions better than a **single-template histogram**.
  - It addresses **intra-patient** and **inter-patient variability**,
  - Accounts for **unimodal** and **bimodal** RR histogram shapes observed in AF [56–58].
  - In contrast, detectors relying on **one merged template** are less effective under high variability conditions.

Alternative Strategy: Comparing Start and End Histograms

- Another method compares **two ΔRR histograms**:
  - One from the **first**, and one from the **last** part of the detection window [35].
  - Instead of matching with template histograms, this method detects **transitions** in rhythm.
- **Detection Parameter**:
  - The **sum of squared differences** between corresponding bin counts of the two histograms.
  - A **small value** indicates **consistent rhythm** (e.g., stable sinus or AF),
  - A **large value** suggests a **rhythm transition** (e.g., sinus to AF or vice versa).
- **Limitations**:
  - The squared difference method alone showed **insufficient performance**.
  - To improve discrimination between **sinus rhythm and AF**, the following were also used:
    - **Number of non-empty bins**
    - **Histogram height**
    - **Standard deviation** of ΔRR intervals

These enhancements help overcome the shortcomings of earlier histogram-based techniques and better capture dynamic rhythm changes.

Shannon Entropy

- **Shannon entropy (ShEn)** measures the **uncertainty or unpredictability** of information content in signals like the **RR interval series** [59].
  - **Higher entropy** → more **uniform** probability distribution (PDF) → **less predictable** signal.
  - **Lower entropy** → more **concentrated** PDF → **more predictable** signal.

Definition:

$$I_{\mathrm{ShEn}} = -\sum_{i=1}^{B} p(x_i)\log_2\left(p(x_i)\right) \qquad (4.7)$$

Where:

- B: number of different values or bins.
- $p(x_i)$: probability of the i-th bin.
- Shannon entropy values range from **0** (fully predictable) to $\log_2(B)$ (fully uniform).
- Often, the entropy value is **normalized** by $\log_2(B)$ for easier interpretation.

Estimation:

- Probabilities $p(x_i)$ are **estimated from histogram counts**:

$$\hat{p}(x_i) = \frac{N(i)}{N} \tag{4.8}$$

- Where:
  - $N(i)$: count of entries in the i-th bin,
  - N: total number of samples.

In the context of AF detection, Shannon entropy helps **quantify RR interval variability**, where **high entropy may indicate AF** due to increased unpredictability.

Shannon Entropy in AF Detection

- **Shannon entropy** I_{ShEn} is typically **larger in AF** than in sinus rhythm, making it useful for AF detection [32].
- **Computation method**:
  - Based on a **modified RR interval series**, where the **longest and shortest RR intervals are removed** to reduce outlier influence.
  - Histogram is constructed from the remaining intervals using **equal-width bins**, spaced between the new minimum and maximum values.
  - At least **16 bins** are recommended for reliable I_{ShEn} estimation.

Limitations at High Heart Rates [49]:

- I_{ShEn} performance **degrades** at heart rates above **90 bpm**.
- **Why?**
  - As heart rate increases, the **probability distribution** $\hat{p}(x_i)$ becomes **narrower**, even if heart rate variation remains constant.
  - Example:
    - At **60 bpm**: 5 bpm variation → RR intervals: 1090 ms to 923 ms → **167 ms spread**.
    - At **120 bpm**: 5 bpm variation → RR intervals: 521 ms to 480 ms → **41 ms spread**.
- Since I_{ShEn} is based on **RR intervals**, not heart rate directly, its power to **distinguish AF from sinus rhythm decreases** as heart rate **increases**.

Symbolic Dynamics and Shannon Entropy

- Instead of computing I_{ShEn} directly from RR intervals, it can be applied to a **symbolic series** derived from **ΔRR intervals** [38].
- This symbolic series is constructed using an **alphabet** (typically 10 symbols), mapped by a **quantization function**.

Symbol Mapping Process:

- Changes in RR intervals are **quantized** by comparing them to a **reference RR series**, which is obtained via **lowpass filtering** of the RR intervals.
- The **quantization grid is dynamic**, adapting to the filtered RR series.
- This approach employs:
  - Linear or time-invariant lowpass filters,
  - Often **ad hoc** in design.
- **Advantage**:
  - Symbolic dynamics helps **improve separation** between **normal beats** and **AF beats** when combined with I_{ShEn}.
  - Likely due to the **quantization step** reducing variability and enhancing signal interpretation.

Follow-up Study [41]:

- Investigated symbolic series vs. Shannon entropy in more detail.
- **Key difference**: This version uses **instantaneous heart rate** (instead of RR intervals) to generate the symbolic sequence.
- A **fixed-step quantization grid** is used.
- **Findings**:
  - Slightly better performance observed with instantaneous heart rate.
  - Although not explicitly explained, this improvement **likely results from bypassing the limitation** seen when using RR intervals at **high heart rates** [49].
- **Conclusion**:

- **Conclusion**:
  - Symbolic dynamics combined with Shannon entropy offers a **robust alternative** to traditional RR interval-based entropy measures,
  - Particularly effective at mitigating **heart rate sensitivity issues** in AF detection.

Sample Entropy

- **Sample entropy** (I_{SampEn}) measures the **self-similarity** or **complexity** of a time series, unlike Shannon entropy which measures symbol probability [60, 61].
  - It is useful for AF detection, as **increased entropy indicates irregularity**, typically found in AF.

Definition:

$$I_{\text{SampEn}} = -\ln\left(\frac{B(m+1,r)}{B(m,r)}\right) \tag{4.9}$$

- $B(m,r)$: probability that pairs of sequences of length m match within tolerance r.
- A **low** I_{SampEn}: signal is **regular**.
- A **high** I_{SampEn}: signal is **irregular**, suggesting AF.

Estimation Procedure:

1. **Segment RR series** $x(0),\ldots,x(N-1)$ into overlapping m**-length vectors**:

$$\mathbf{x}(i) = \begin{bmatrix} x(i) \\ \vdots \\ x(i+m-1) \end{bmatrix}, \quad i = 0,\ldots,N-m-1 \tag{4.10}$$

2. **Compare vectors** $\mathbf{x}(i)$ and $\mathbf{x}(j)$ using **maximum norm**:

$$\|\mathbf{x}(i)-\mathbf{x}(j)\|_\infty = \max_{k=0,\ldots,m-1} |x(i+k)-x(j+k)| \tag{4.11}$$

3. Two sequences are **similar** if:

$$\|\mathbf{x}(i)-\mathbf{x}(j)\|_\infty \le r$$

4. Compute **average number of similar subsequences** (excluding self-matches):

$$\hat{B}(m,r) = \frac{1}{N-m-1}\sum_{\substack{j=0\\ j\neq i}}^{N-m-1} H\left(r - \|\mathbf{x}(i)-\mathbf{x}(j)\|_\infty\right) \tag{4.12}$$

- $H(z)$: **Heaviside step function**:

$$H(z) = \begin{cases} 1, & z \ge 0 \\ 0, & z < 0 \end{cases} \tag{4.13}$$

In Summary:

- **Regular signals (sinus rhythm) → low** I_{SampEn}
- **Irregular signals (AF) → high** I_{SampEn}
- Sample entropy effectively captures **transitions** from regular to irregular rhythm.

Estimation of Similarity Probability for Sample Entropy

- The probability of two m-length subsequences being similar is estimated by:

$$\hat{B}(m,r) = \frac{1}{N-m}\sum_{i=0}^{N-m-1} \hat{B}_i(m,r)$$

- Expanded form:

$$\hat{B}(m,r) = \frac{1}{(N-m)(N-m-1)}\sum_{i=0}^{N-m-1}\sum_{\substack{j=0\\ j\neq i}}^{N-m-1} H\left(r - \|\mathbf{x}(i)-\mathbf{x}(j)\|_\infty\right) \tag{4.14}$$

Where:

- $\hat{B}_i(m,r)$ is the proportion of subsequences similar to $\mathbf{x}(i)$,
- $H(\cdot)$ is the **Heaviside function**, returning 1 if the argument is ≥ 0, 0 otherwise.

Handling Issues in Short Detection Windows:

- For **short windows**, which are common in **brief AF episode detection**, the risk arises that **no matches are found**, especially with small r.
  - This leads to $\hat{B}(m,r) = 0$, making I_{SampEn} **undefined**.
- **Solution**:
  - Convert probabilities to **densities** by dividing by the volume of the matching region [62]:

$$-\ln\left(\frac{B(m+1,r)}{(2r)^{m+1}}\right) + \ln\left(\frac{B(m,r)}{(2r)^m}\right) = -\ln\left(\frac{B(m+1,r)}{B(m,r)}\right) + \ln(2r) \tag{4.15}$$

- This **normalization** allows comparison across different r values and reduces dependence on scale.

Practical Implications:

- r is usually set as a **fraction of the standard deviation** of the RR interval data [60],
- But it can also be **dynamically increased** until $\hat{B}(m,r) > 0$, especially in **AF analysis**, where the likelihood of matches can be low.

Refinements of Sample Entropy for AF Detection

- **Coefficient of Sample Entropy (CSampEn)** enhances regular sample entropy by incorporating the **mean RR interval length** $\hat{m}_x$, improving detection **independent of heart rate** [34].

$$I_{\text{CSampEn}} = I_{\text{SampEn}} + \ln(2r) - \ln(\hat{m}_x) \tag{4.16}$$

- This modification ensures:
  - **Higher CSampEn** in **AF** (where heart rate is higher),
  - **Lower CSampEn** in **sinus rhythm** (where heart rate is lower).

Subsequence Length Choice (m):

- Shortest possible subsequence, $m = 1$, is often chosen because:
  - **Autocorrelation** in RR intervals during AF is nearly **zero** [3],
  - Empirical results show **better performance** at $m = 1$ than at larger values [34],
  - Further discussion on selecting m and r is found in **Sect. 6.4.4**.

Alternative Entropy Measures:

1. **Approximate Entropy (ApEn)** [63]:
   - Similar to SampEn, but **includes self-matches** in (4.12).
   - Has known limitations:
     - **Bias** due to sample size,
     - Lack of **relative consistency** [60],
     - Therefore, **less commonly used** than SampEn.
2. **Fuzzy Entropy**:
   - Replaces the **Heaviside function** $H(z)$ with a **fuzzy function** to allow **gradual similarity scoring** [64].
   - Applied in:
     - **Heart rate variability** studies [65],
     - **f wave analysis** [66].

- While promising, it is still under investigation whether fuzzy entropy **outperforms SampEn** in AF detection tasks.

Probability of Pairs of Matching RR Interval Subsequences

- A simplified approach for AF detection is to compute only the probability $B(m,r)$, part of the definition of sample entropy I_{SampEn} [40, 67].
  - **Advantages**:
    - **No need** for $B(m+1,r)$,
    - Avoids computing logarithms or ratios,
    - **Avoids undefined values** when no matches are found.

Revised Estimator $\hat{C}(m,r)$:

- Instead of using the **maximum norm** from (4.12), this estimator uses the **Euclidean norm**:

$$\hat{C}(m,r) = \frac{2}{(N-m)(N-m-1)} \sum_{i=0}^{N-m-1} \sum_{j=i+1}^{N-m} H(r - \|\mathbf{x}(i) - \mathbf{x}(j)\|) \qquad (4.17)$$

- Where:
  - $\|\cdot\|$: Euclidean norm,
  - $H(\cdot)$: Heaviside function,
  - **Self-matches excluded**, and **normalization** uses the total number of unique pairs.
- This estimator is similar to $\hat{B}(m,r)$ but:
  - **Computes pairwise distances only once**,
  - Reduces computational complexity.

Simplified AF Detector: $B(m=1,r)$

- A very efficient approach uses **only scalar differences** between RR intervals:

$$\hat{B}(m=1,r) = \frac{2}{(N-1)(N-2)} \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} H(r - |x(i) - x(j)|) \qquad (4.18)$$

- This version:
  - Avoids constructing subsequences,
  - Bypasses both max and Euclidean norms,
  - Makes **real-time AF detection more feasible**.

These estimators allow entropy-inspired analysis while avoiding the computational burdens of traditional sample entropy, especially useful in **short windows** or **low-power devices**.

Simplified Sample Entropy (SSampEn)

- Before thresholding, the probability $\hat{B}(m=1,r)$ (see Eq. 4.18) is normalized by the **mean RR interval length** in the detection window.
- This leads to a new parameter:
  Simplified Sample Entropy, denoted I_{SSampEn} [40]:

$$I_{\text{SSampEn}} = \frac{\hat{B}(m=1,r)}{\tilde{m}_x} \qquad (4.19)$$

Key Elements:

- $\tilde{m}_x$: **Mean RR interval**, computed using **exponential averaging**, excluding ectopic beat intervals (see Sect. 4.2.5).
- Purpose: Reflect that **AF is typically accompanied by increased heart rate**, i.e., **lower RR intervals**.

Relation to Other Measures:

- Structurally similar to the **coefficient of variation** (Eq. 4.2):
  - **Numerator**: a dispersion measure (though threshold-based),
  - **Denominator**: mean of RR intervals (same as in Eq. 4.2).

Heart Rate Adaptation of r:

- An alternative to fixed r is to define it as a **function of the heart rate**:

$$r \to r(\bar{m}_x)$$

  - This adjusts sensitivity dynamically with respect to heart rate.
  - If a fixed r is still used, it can be defined as a **fraction of the standard deviation** over a large dataset [60].

4.2.2 Poincaré-Based Parameters

- The **Poincaré plot** visualizes **pairs of successive RR intervals**, $(x(n), x(n+1))$, and is used to **characterize cardiac rhythms**.
  - Originally applied in **heart rate variability** studies over long recordings (up to several days) [69–72].
  - In AF detection, used with **shorter detection windows** (typically 60–120 s).
- **AF detection relevance**:
  - AF produces a **highly scattered Poincaré plot**, compared to more concentrated patterns from **normal sinus rhythm** or **ectopic beats** (illustrated in Fig. 4.4).

Main challenge:

- Translating **scattering patterns** into **quantifiable detection parameters**.

Two main approaches:

1. **Density-based parameters**:
   - Reflect how densely points populate different regions of the plot [33, 51, 73].
2. **Geometric-based parameters**:
   - Analyze the **shape and distribution** of points [50].

Variations in plotting axes:

- Instead of just $(x(n), x(n+1))$, other pairings can be used:
  - $(\Delta x(n), \Delta x(n+1))$
  - $(x(n), \Delta x(n))$
- These variations also reflect **beat-to-beat irregularity**, and may offer better sensitivity to **AF patterns**.

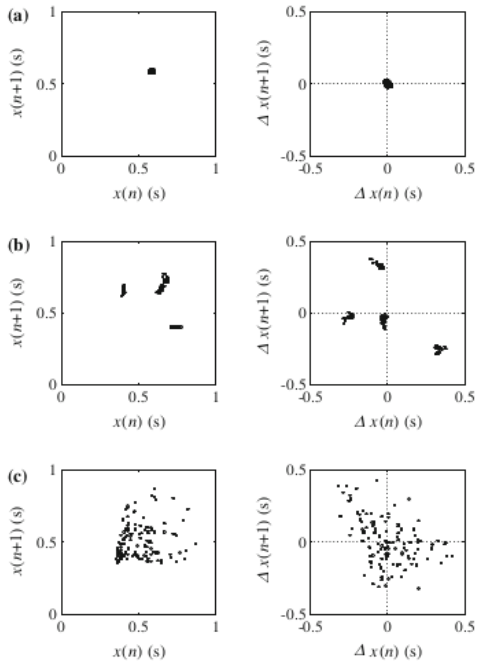


Fig. 4.4 – Poincaré Plots in RR Interval Analysis

- **Poincaré plots** are shown for three rhythm conditions:
  - Left column: $(x(n), x(n+1))$
  - Right column: $(\Delta x(n), \Delta x(n+1))$
  - All based on 128 RR intervals

(a) Normal Sinus Rhythm

- **Left (x, x+1)**:
  - Clustered near the diagonal → very regular pattern
- **Right (Δx, Δx+1)**:
  - Tight cluster around (0, 0) → minimal beat-to-beat variability

---

(b) Sinus Rhythm with Ectopic Beats

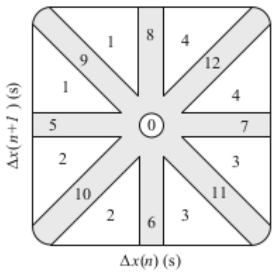
- **Left (x, x+1)**:
  - Multiple clusters → reflects irregular insertions due to ectopic beats
- **Right (Δx, Δx+1)**:
  - Distinct scattered regions in multiple quadrants → sign of pattern breaks

(c) Atrial Fibrillation (AF)

- **Left (x, x+1)**:
  - Widely scattered across the plot → high irregularity
- **Right (Δx, Δx+1)**:
  - Broad, diffuse spread across all quadrants → strong beat-to-beat variability

Key Takeaways:

- **AF detection via Poincaré plot**:
  - The plot based on $(\Delta x(n), \Delta x(n+1))$ offers better sensitivity to irregularity than $(x(n), x(n+1))$.
  - It captures **positive and negative changes**, allowing **all four quadrants** to be analyzed.
- **Binning strategy**:
  - The scatter area is divided into **small square cells**, treated as 2D histogram bins.
  - Recommended **bin size** is **25 ms** for precise characterization [51, 76].



Advanced Use of Poincaré Plots in AF Detection

- The **Poincaré plot** defined by $(\Delta x(n), \Delta x(n+1))$ is divided into **12 regions** (see Fig. 4.5), allowing the **population of each region** to be correlated with **different rhythm types**.

Detection Steps:

1. **Count "nonzero bins"**:
   - Bins (cells) with at least one point are tallied across all regions.
   - Region 0 (central circular zone) is typically populated by **normal sinus rhythm**.
2. **Apply corrections**:
   - Subtract:
     - Points in **region 0**,
     - A factor accounting for **atrial premature beats (APBs)**, which tend to cluster near region 0.
3. **Detection rule**:
   - An **AF episode** is flagged when the **corrected total number of nonzero bins** exceeds a fixed threshold,
   - But only if the number of points in regions associated with **atrial tachycardia** falls **below another threshold**.

Region-based rhythm classification (from Fig. 4.5):

- **Region 0**: Normal sinus rhythm.
- **Regions 6, 7, 9, 11**: Atrial tachycardia.
- **Regions 1–4**: Atrial and ventricular premature beats.
- **All other regions**: Associated with AF.

Simplified Non-Region-Based Method:

- Using only the **total number of nonzero bins** (parameter P_{NZPP}) provides a **simpler detection rule**:
  - All bins in the plot space (grid of $\Delta x(n)$ vs. $\Delta x(n+1)$) are treated equally.
  - An AF episode is detected when the count **exceeds a threshold**.
  - This method **avoids region segmentation**, making it ideal for **real-time applications** [33].

Use Cases:

- **Region-based method**: More accurate classification (e.g., distinguishes AF from tachycardia).
- **Non-region-based (grid-only)**: Easier to implement, especially in **implantable or portable devices**.

This dual-method approach to Poincaré analysis balances **detection accuracy** with **computational efficiency**.

Geometrical Characterization in Poincaré-Based AF Detection

- **Limitation of histogram-based detectors**:
  - Require many RR intervals for sufficient bin population in 2D histograms.

- A **2-minute detection window** is often needed [51], though **shorter windows (e.g., 64 beats)** are possible with region-based or grid-only methods [33].
- **Advantage of Poincaré plot using** $(\Delta x(n), \Delta x(n+1))$:
  - Better detection of AF and other rhythms (e.g., **APBs**), even with **shorter data segments**.
  - The benefit of using this form over $(x(n), x(n+1))$ is still under investigation.

Second Approach: Geometrical Poincaré Analysis

- Focuses on **how** the points populate the Poincaré plot, rather than **where**.
- Conceptually related to **statistical dispersion**, but uses **geometric spread** as the quantification metric [50, 77].

Elliptical Shape Analysis (in Sinus Rhythm)

- In sinus rhythm:
  - Points in $(x(n), x(n+1))$ plot align along the **identity line** $x(n) = x(n+1)$, forming an **elliptical cloud**.
- To analyze dispersion:
  - Apply a **45° rotation** to obtain axes aligned with and perpendicular to the line of identity:

$$\begin{bmatrix} y(n+1) \\ y(n) \end{bmatrix} = \begin{bmatrix} \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \\ \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \end{bmatrix} \begin{bmatrix} x(n+1) \\ x(n) \end{bmatrix}, \quad n = 0, \ldots, N-2 \qquad (4.20)$$

- This transforms the data such that:
  - $y(n+1)$: along the **line of identity**,
  - $y(n)$: **perpendicular** to it.

Quantifying Ellipse Shape

- Use standard deviations along both axes:

$$\sigma_y = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-2} (y(n+j) - \tilde{m}_y)^2}, \quad j = 0, 1 \qquad (4.21)$$

- Where:
  - $\tilde{m}_y$: mean of the transformed data,
  - $\sigma_{y,1}$: dispersion along line of identity,
  - σ_y: dispersion perpendicular to it.
- In AF, the **cloud becomes more circular**, increasing σ_y relative to $\sigma_{y,1}$, reflecting **higher RR variability**.

Conclusion:
Geometrical Poincaré parameters offer a **continuous** way to measure RR variability, especially useful in AF detection where **scatter increases** and symmetry shifts from elliptical (sinus) to circular (AF).

AF Detection Using Poincaré Plot Dispersion

- In **AF**, the RR interval pattern deviates significantly from **sinus rhythm**:
  - The **elliptic cluster** assumption in sinus rhythm **no longer holds**.
  - Therefore, while σ_y (perpendicular dispersion) is still useful [50, 79], its counterpart $\sigma_{y,1}$ becomes **less meaningful**.

Transformation for AF Analysis

- The transformation in Eq. (4.20) leads to a simplified expression for AF:

$$y(n) = \frac{1}{\sqrt{2}}(x(n+1) - x(n)) = \frac{\Delta x(n)}{\sqrt{2}} \qquad (4.22)$$

- Since successive RR intervals are **highly irregular** in AF:
  - The **mean value** of $y(n)$ is approximately **zero**,
  - Making the **standard deviation** $\sigma_{y,0}$ an effective measure of **dispersion** along the diagonal.

Approximated Dispersion Formula:

$$\sigma_{y,0} \approx \sqrt{\frac{1}{2(N-1)} \sum_{n=1}^{N-1} \Delta x^2(n)} \qquad (4.23)$$

- This is mathematically identical to the **RMSSD measure** from earlier (Eq. 4.3), but derived **within the Poincaré plot framework**.

Key Insights:

- **Dispersion along the identity line**, $\sigma_{y,0}$, is a direct representation of **beat-to-beat variability**.
- When using **short-term ECG data**:
  - The Poincaré plot loses its **elliptical shape**,
  - It begins to resemble **dot-like** scatter in AF (as seen in Fig. 4.4),
  - Therefore, **ellipse-based analysis** becomes less reliable for AF detection in short windows.

Conclusion:

- While **ellipse-fitting** works well in **long-term ECG**, for **short-term AF detection**, simpler measures like **RMSSD** or $\sigma_{y,0}$ derived from **ΔRR intervals** are more effective and interpretable within a geometric framework.

Geometrical RR Dispersion Parameters from the Poincaré Plot

- Another **geometrical detection parameter**, denoted σ_c, is based on the **Euclidean distance** between:
  - $(x(n), x(n+1))$ and
  - $(x(n+1), x(n+2))$
    – measuring **local rate of change** in RR intervals [50].

Definition of σ_c:

$$\sigma_c = \frac{1}{N-2}\sum_{n=1}^{N-2}\sqrt{\Delta x^2(n) + \Delta x^2(n+1)} \tag{4.24}$$

- Can be equivalently written as:

$$\sigma_c = \frac{1}{N-2}\sum_{n=1}^{N-2}\sum_{k=0}^{1}\Delta x^2(n+k) \tag{4.25}$$

- Like $\sigma_{y,0}$, it represents a measure of **RR interval dispersion**, especially the **local variation** between consecutive RR pairs.

Normalization by Mean RR Length:

- For consistent interpretation, both $\sigma_{y,0}$ and σ_c are **normalized** by the **mean RR interval** $\tilde{m}_x$:

$$\sigma'_{y,0} = \frac{\sigma_{y,0}}{\tilde{m}_x} \tag{4.26}$$

- This aligns with:
  - **Coefficient of sample entropy** (Eq. 4.16),
  - **Simplified sample entropy** (Eq. 4.19).

Purpose of Normalization:

- Ensures that **higher heart rates** (i.e., **shorter RR intervals**) lead to **increased normalized dispersion**,
- Enhancing the ability to **discriminate AF**, which is often accompanied by **both higher heart rate and greater irregularity**.

Conclusion:
Parameters σ_c and $\sigma_{y,0}$, especially in their **normalized forms**, are efficient and interpretable geometrical tools for **RR irregularity assessment**, aiding in **robust AF detection**.

4.2.3 Time-Varying Coherence Function

- This method uses a **linear systems approach** to detect AF by analyzing the **spectral coherence** of RR intervals in two adjacent windows [37].
- **Key Concept**:
  - During **sinus rhythm**, spectral coherence remains high and stable.
  - At the **onset or end of AF**, coherence drops abruptly.

Definition of Time-Varying Coherence Function (TVCF)

- Let $x(n)$ and $y(n)$ be two successive windows of RR data (input/output of a linear system):

$$C_{xy}(\omega, n) = \frac{|S_{xy}(\omega, n)|^2}{S_x(\omega, n)S_y(\omega, n)} \tag{4.27}$$

$$C_{yx}(\omega, n) = \frac{|S_{yx}(\omega, n)|^2}{S_y(\omega, n)S_x(\omega, n)} \tag{4.28}$$

- Where:
  - $S_{xy}(\omega, n)$: **Time-varying cross-spectrum**,
  - $S_x(\omega, n), S_y(\omega, n)$: **Time-varying spectra** of $x(n)$ and $y(n)$.

Overall TVCF Measure:

- To consider both directions (x → y and y → x), use:

$$C^2(\omega, n) = C_{xy}(\omega, n)C_{yx}(\omega, n) \tag{4.29}$$

Time-Varying Transfer Functions:

- When viewing either signal as input:

$$H_{x\to y}(\omega, n) = \frac{S_{xy}(\omega, n)}{S_{x}(\omega, n)} \tag{4.30}$$

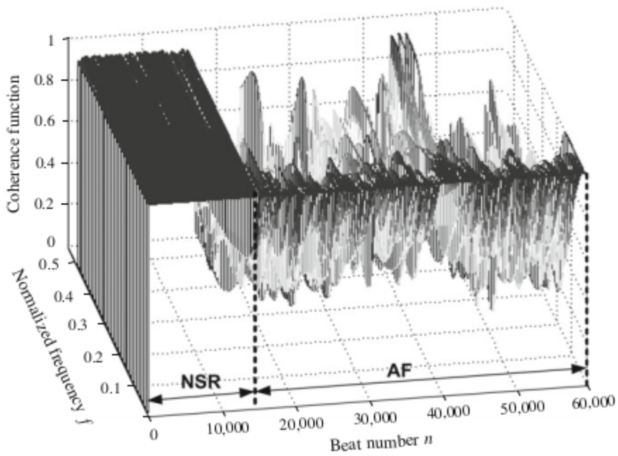
$$H_{y\to x}(\omega, n) = \frac{S_{yx}(\omega, n)}{S_y(\omega, n)} \quad (4.31)$$

- So, the TVCF can also be expressed as:

$$C^2(\omega, n) = |H_{x\to y}(\omega, n) H_{y\to x}(\omega, n)|^2 \quad (4.32)$$

Application in AF Detection:

- When $C^2(\omega, n)$ **drops**, it indicates a **loss of coherence**—a likely sign of **AF onset or termination**.
- This method detects AF by **monitoring coherence changes** over time, leveraging **frequency-domain analysis** of RR series.



Model-Based Estimation of Time-Varying Coherence Function

- The transfer functions $H_{x\to y}(\omega, n)$ and $H_{y\to x}(\omega, n)$ are estimated using a **model-based approach**:
  - Assumes both windows follow an **ARMA model** (autoregressive moving average) [81].
  - **Preferred over spectrogram-based methods** due to better **frequency resolution**, if the model fits well.
- **Model tuning**:
  - Parameters and order are optimized using a specialized technique.
  - **Longer detection windows** → require **higher model orders** for accurate TVCF estimation.

Fig. 4.6: Behavior of $C^2(\omega, n)$ in AF vs. NSR

- Shows $C^2(\omega, n)$ computed over an RR interval series containing a transition from:
  - **Normal sinus rhythm (NSR) → Atrial fibrillation (AF)**
  - Detection windows: 128 beats, sliding by 128 beats per step.

Observations:

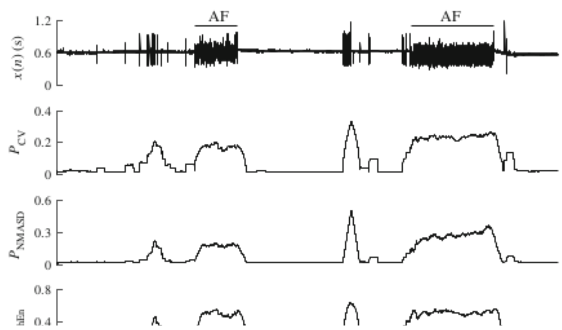
- **During NSR**:
  - $C^2(\omega, n)$ is **high** and **uniform** across frequencies.
  - Implies high coherence between adjacent RR segments.
- **At AF onset**:
  - **Coherence drops abruptly**,
  - **Variation across frequency increases**, especially at **higher frequencies**.

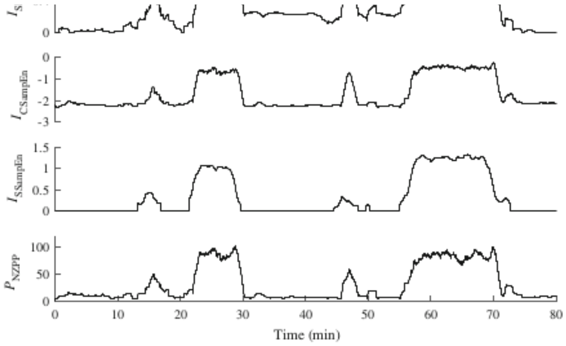
Detection Strategy:

- Use the **variance of** $C^2(\omega, n)$ across frequency for each beat n,
- Higher variance → indicates a **transition to AF**,
- This variance is used as a **detection parameter** for identifying AF onset.

Conclusion:

By leveraging ARMA models and the time-varying coherence function, this method provides a **sensitive frequency-domain metric** for detecting **dynamic changes in rhythm**, especially the transition from **regularity (NSR)** to **irregularity (AF)**.





4.2.4 Parameter Time Series Exemplified

- An 80-minute ambulatory ECG with **two AF episodes** and several **ectopic beat runs** was analyzed.
- **Detection parameters** were computed using:
  - A **128-beat sliding detection window** (sliding by one beat),
  - Exception: I_{SampEn}, which used an **8-beat window** [40].

Key Observations from Fig. 4.7:

1. **Clear distinction** between **normal sinus rhythm** and **AF episodes** across all detection parameters.
2. **Impact of ectopic beats**:
   - Strongest for:
     - P_{CV}
     - P_{NMASD}
   - These parameters spike during ectopic runs—**even outside AF episodes**.
   - In contrast, I_{SSampEn} is **less affected**, showing robustness.

   ➡ This indicates the need to **mitigate ectopic influence** in AF detection (see Sect. 4.2.5).
3. **Background fluctuations**:
   - I_{ShEn} shows **more pronounced fluctuations** during **sinus rhythm**, suggesting higher **baseline variability** compared to other parameters.

Conclusion:
This section demonstrates how different parameters behave over time and under various cardiac events. It highlights the **robustness of some measures** (like I_{SSampEn}) and the **sensitivity of others to ectopic beats**, which must be considered when designing **reliable AF detectors**.

Handling Ectopic Beats in Rhythm-Based AF Detection

- **Ectopic beats** (VPBs and APBs) are common and **can trigger false AF detections**, especially in rhythm-based detectors.
- Properly **excluding or flagging** RR intervals linked to ectopic beats is critical:
  - It **increases specificity**,
  - But must be done carefully to **preserve sensitivity** to actual AF episodes.

Current Practices and Challenges

- Many detectors **do not implement ectopic beat filtering** explicitly.
- When the **ΔRR interval histogram** is used for detection (e.g., with the **Kolmogorov–Smirnov test**) [31], rhythms with frequent VPBs may be **misclassified as AF**.
- The problem arises from the **compensatory pause**:
  - Negative ΔRR followed by positive ΔRR mimics AF.
  - This creates a **shoulder** in the histogram around **400–600 ms**—similar to that seen in AF histograms.

Fig. 4.7 – Time Series Comparison of Detection Parameters

- **Top panel**: RR interval series $x(n)$ with two annotated AF episodes.
- **Detection parameters plotted**:
  1. P_{CV}: Coefficient of variation
  2. P_{NMASD}: Normalized mean absolute successive differences
  3. I_{ShEn}: Shannon entropy
  4. I_{CSampEn}: Coefficient of sample entropy
  5. I_{SSampEn}: Simplified sample entropy
  6. P_{NZPP}: Number of nonzero Poincaré bins

Insights from Fig. 4.7:

- All parameters clearly detect both **AF episodes**.
- **Ectopic beats before the second AF episode** cause:
  - Minimal change in I_{SSampEn},
  - Pronounced spikes in P_{CV} and P_{NMASD},
  - Highlighting their **sensitivity to ectopic beats**.
- I_{ShEn} shows **more background fluctuation** even during sinus rhythm.

Mitigation Strategy (Preliminary):

- **Reducing false alarms** from VPBs:
  - Possible through analyzing **shoulder height and width** in the ΔRR histogram.
  - However, no formal method exists yet to **detect or quantify** such a shoulder [31].

Conclusion:
Robust AF detection requires **handling ectopic beat influence**, particularly in histogram- and entropy-based methods. Among tested parameters, I_{SSampEn} and I_{CSampEn} show **greater robustness**, making them strong candidates for reliable, real-world AF detection.

Ectopic Beat Identification Using Poincaré Plots and Ratio Conditions

- When using **Poincaré plots** for AF detection, **ectopic beats** can be spotted by their **distinct clustering** in specific bins:
  - **Bigeminy**: tightly clustered points in a few bins [33, 51].
  - **AF**: widely scattered points.
  - **Problem**: when using $(x(n), x(n+1))$, variable heart rate can **smear ectopic clusters**, increasing **false alarms**.
  - **Solution**: using $(\Delta x(n), \Delta x(n+1))$ reduces this effect by focusing on beat-to-beat variability.

RR Ratio-Based Ectopic Beat Filtering ([32], [37])

To detect and remove **ventricular premature beats (VPBs)** from the RR series, a method based on **three RR interval ratios** is used:

An RR interval $x(n)$, and its surrounding intervals $x(n-1), x(n+1), x(n+2)$, must satisfy all of the following to be classified as a VPB and excluded:

1. **Preceding short RR** (VPB trigger):

$$\frac{x(n)}{x(n-1)} < \gamma_1 \tag{4.33}$$

2. **Compensatory pause**:

$$\frac{x(n)}{x(n+1)} > \gamma_{99} \tag{4.34}$$

3. **Ratio of next two beats**:

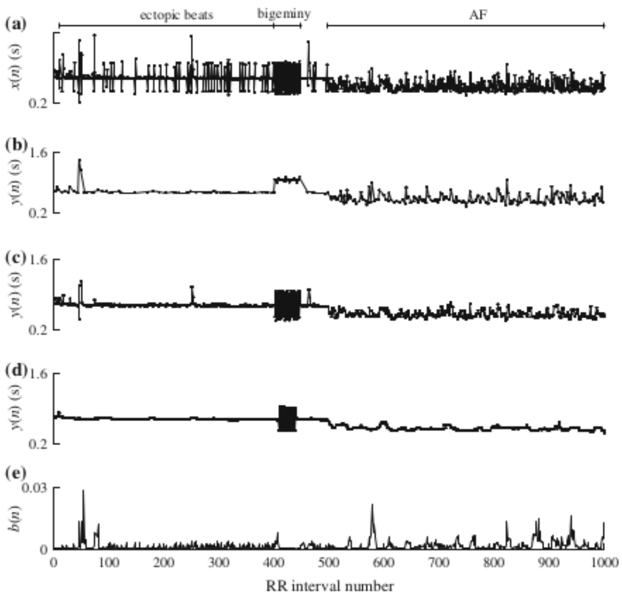
$$\frac{x(n+1)}{x(n+2)} > \gamma_{25} \tag{4.35}$$

- Thresholds:
  - $\gamma_1, \gamma_{25}, \gamma_{99}$:
    1st, 25th, and 99th **percentiles** of the RR interval ratio distribution.

Practical Considerations:

- Thresholds adapt to the **current ECG segment**,
- As the Poincaré plot and RR ratios reflect **local irregularity**, these rules offer **automated VPB exclusion**,
- Visual examples are referenced in Fig. 4.8a and 4.8b.

Conclusion:
These RR ratio-based rules are an effective way to **exclude ectopic beats** without affecting genuine AF intervals, helping to **reduce false positives** in rhythm-based AF detectors.



Median Filtering and Ad Hoc Tests for Ectopic Beat Removal

- **Median filtering** is commonly used to suppress ectopic beats in RR interval sequences:
  - It preserves **sharp transitions** at the **onset/offset of AF episodes**,
  - Filters of **lengths 3 to 17** have been used:
    - **Longer filters** remove ectopic beats more effectively,
    - But may **miss short AF episodes**—making **shorter filters** preferable when detecting **brief AF**.

Fig. 4.8 Summary:

- Shows an RR interval series $x(n)$ with:
  - **Ectopic beats**, **bigeminy**, and an **AF episode** starting around sample 500.

Panels:

- **(a)**: Original $x(n)$.
- **(b)**: Output after applying the **three ratio-based ectopic filtering conditions** (Eqs. 4.33–4.35).
- **(c)**: Output after **3-point median filtering**.
- **(d)**: Output after **17-point median filtering**—more effective at flattening ectopic patterns and clarifying AF.
- **(e)**: Function $b(n)$ (Eq. 4.36, not shown here), used to **flag bigeminy**, not replace RR intervals.

Ad Hoc Ratio Tests with Pattern Matching

- An extension of Eqs. 4.33–4.35:
  - Detects common **non-AF arrhythmias** like bigeminy/trigeminy,
  - Builds a **template database** of RR ratio sequences.
  - During detection:
    - The RR ratio sequence in the detection window is **compared to templates**,
    - If matched, **AF is ruled out**.
- This approach allows **classification of rhythm type** before computing AF detection parameters.

Notes on Implementation:

- This method integrates **ectopic filtering directly into the classifier**,
- Several **thresholds** must be set (e.g., matching criteria, filtering length),
- However, **performance sensitivity** to these thresholds **has not yet been well established**.

Conclusion:
Median filtering and **pattern-based ad hoc tests** significantly improve detection performance by reducing **false AF alarms** caused by ectopic beats—especially **bigeminy**. Balancing **filter length** and **sensitivity to AF transitions** remains a key consideration.

Flag Function for AF Detection

A **simple flag function** $b(n)$ has been introduced [40] to **indicate whether the rhythm is likely AF**, defined as:

$$b(n) = \left(\frac{\sum_{m=0}^{M-1} x_m(n-m)}{\sum_{m=0}^{M-1} x(n-m)} - 1 \right)^2, \quad n = M, \ldots, N-1 \qquad (4.36)$$

- Where:
  - $x(n)$: original RR interval,
  - $x_m(n)$: output of a **3-point median filter**,
  - M: even integer specifying window length,
  - N: total number of RR intervals.

Interpretation:

- For **regular rhythms** and **bigeminy**:
  - Median-filtered signal $x_m(n)$ ≈ original signal $x(n)$,
  - → ratio ≈ 1 → $b(n) \approx 0$ (no AF).
- For **AF rhythms**:
  - Median filtering smooths the signal (removes short-term irregularity),
  - So $x_m(n)$ deviates from $x(n)$,
  - → $b(n)$ becomes **larger**, flagging **likely AF**.

Why the Square?

- The squaring operation:
  - Emphasizes differences,
  - Ensures **positive output**,
  - Improves separation between **AF** and **non-AF** rhythms.

Application:

- Used as a **weighting function**, not to exclude RR intervals.
- Helps **highlight AF onset** while tolerating ectopic beats (e.g., bigeminy).
- Refer to **Fig. 4.8e**:
  - $b(n)$ stays near **0** during bigeminy,
  - **Rises sharply** during AF.

Conclusion:
The $b(n)$ flag is a lightweight, signal-based metric that enhances **real-time AF detection**, especially useful in **implantable monitors** and **mHealth devices**. Continued development of such tools is needed to refine **ectopic beat handling** in practical applications.

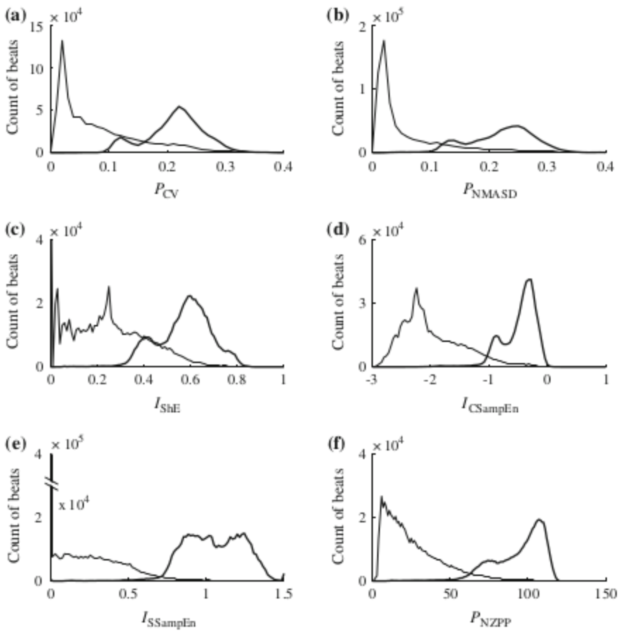
4.2.6 Classification

Threshold-Based Classification:

- The most common method for AF classification is to apply **thresholds** to one or more **features** derived from RR interval irregularity:
  - Could be a **single feature** [31, 33, 34, 38, 40],
  - Or **multiple features** with their own thresholds [35].
- **Threshold determination**:
  - Often optimized using a **performance metric**, e.g., **area under the ROC curve** (see Sect. 4.5),
  - Can also be determined via **statistical assumptions** [31].

Feature Correlation & Selection:

- In multi-feature classifiers, it's critical to ensure features are not **redundant**:
  - Use **Principal Component Analysis (PCA)** to reduce correlated features and retain only the most **informative**,
  - Reduces dimensionality and improves **robustness**, especially when training data is limited [82].
- **Feature selection in AF detection**:
  - Methods like **sequential forward floating selection** have been used [46],
  - But not yet widely adopted across AF detection systems.



Feature Relevance & Simplicity:

- Important to assess **feature importance** for low-power, real-time devices.
- No current studies systematically quantify **individual feature contributions**, but initial results highlight the dominant role of **rhythm irregularity** over P-wave/f-wave morphology [83].

Histogram Overlap Method (Feature Comparison):

- A practical method to evaluate feature effectiveness:
  - Compare **histograms of RR-based features** between **AF and non-AF windows** using the **same dataset** [38, 40, 41, 49],
  - The **less overlap** → the more **discriminative** the feature.
- **Fig. 4.9** (referenced here):
  - Based on 128-beat detection windows (except SampEn: 8 beats),
  - I_{SSampEn} yields the **least overlap → most suitable** feature,
  - **Number of nonzero bins** in the Poincaré plot and I_{CSampEn} also rank high,
  - **Shannon entropy** (I_{ShEn}) has **greatest overlap**, suggesting **lower effectiveness**.

Conclusion:

Conclusion:
Effective AF classifiers prioritize **low-dimensional**, **uncorrelated**, and **highly discriminative features**. Current evidence suggests that features representing **rhythm irregularity**, especially I_{SSampEn}, are most effective. Further refinement in **feature selection and analysis** is essential for real-time, resource-limited applications.

Pattern Classification Approaches in AF Detection

Beyond traditional threshold-based classifiers, more advanced **pattern classification techniques** have been explored for AF detection:

---

1. Support Vector Machines (SVMs)

- References: [39, 50, 84]
- Advantages:
  - Can model **nonlinear decision boundaries** [85],
  - Require only **two design parameters**, both related to **penalizing misclassifications**.

---

2. Linear Discriminant Analysis (LDA)

- Reference: [36]
- Characteristics:
  - Simpler than nonlinear models, but requires:
    - **Mean vectors** and **covariance matrices** for both AF and non-AF data,
    - More **computationally intensive** during training than threshold tests.

Feature Vector Dimension

- In these studies, feature vectors ranged from **2 to 24 dimensions**.
- Trade-off:
  - Higher dimensions → potentially **better discrimination**,
  - But also lead to **increased complexity** and risk of **overfitting** with small datasets.

Conclusion:
SVMs offer **flexibility and simplicity** in tuning, while LDA provides a **statistical approach** at the cost of **higher training effort**. These methods represent an evolution beyond simple thresholds toward more **robust, adaptive AF classifiers**.

Performance and Feature Distributions in AF Detection (Fig. 4.9)

Classifier Performance Insights (from Table 4.1)

- Surprisingly, a **single-threshold detector** [40] outperformed an **SVM-based classifier** [39].
- This may seem counterintuitive since:
  - SVMs offer **flexible, nonlinear decision boundaries**.
- Possible reasons:
  - **SVM did not generalize well** due to a **small or unrepresentative training set**.
  - **Weaker feature selection** may have hindered SVM performance.

Fig. 4.9 – Feature Distributions: AF vs Non-AF

Histograms for six detection parameters are shown:

- **Thick line**: AF
- **Thin line**: non-AF

(a) P_{CV} – Coefficient of Variation

- Substantial overlap between AF and non-AF,
- Indicates **limited discriminative power**.

(b) P_{NMASD} – Normalized Mean of Absolute Successive Differences

- Strong peak for AF,
- But still noticeable overlap.

(c) I_{ShEn} – Shannon Entropy

- Very large overlap,
- **Least effective** feature.

(d) I_{CSampEn} – Coefficient of Sample Entropy

- Better separation between classes,
- More useful for AF detection.

(e) I_{SSampEn} – Simplified Sample Entropy

- **Best separation** observed,
- Strong candidate for robust detection.

(f) P_{NZPP} – Nonzero Poincaré Bins

- Clear separation,
- Especially effective for capturing irregular rhythms.

Conclusion:

- **Simple threshold detectors** can outperform more complex classifiers (like SVMs) if **well-chosen features** are used.

- I_{SSampEn} and P_{NZPP} provide the **clearest separation** between AF and non-AF.
- Detection performance hinges more on **feature quality** than on the complexity of the classifier itself.

Machine Learning Classifiers vs Threshold-Based AF Detection

Current Status of Machine Learning in AF Detection:

- **Machine learning approaches** (e.g., SVMs) have **not yet outperformed** traditional **threshold-based** methods in AF detection.
- One key reason: many classifiers **lack robustness** to **non-AF arrhythmias**, such as:
  - **Bigeminy and trigeminy**
  - **APBs and VPBs**
  - **Supraventricular tachycardia**
  - **Atrioventricular junctional rhythms**
- These **confounding rhythms** often mimic AF patterns and increase **false positives** when not handled properly.

Importance of Ectopic Beat Handling Before Classification

- Handling ectopic beats **before classification** can significantly **boost performance**.
- Example from [40] (unpublished result):
  - Detector uses:
    - I_{SSampEn}: Simplified sample entropy (8-beat window),
    - $b(n)$: Flag function indicating AF-likelihood (see Eq. 4.36),
  - These are **fused** and passed through a **simple threshold**.

Performance Gain (Using NSRDB Dataset)

- NSRDB includes frequent bigeminy (see Sect. 3.1).
- Incorporating $b(n)$ into detection led to:
  - **Specificity increase**: from **93.2% → 98.6%**
  - **Sensitivity remained unchanged**
- Demonstrates the value of handling ectopic beats **before applying classifiers**.

Conclusion:
While machine learning has potential, **threshold-based detectors remain superior**—especially when **robust ectopic beat filtering** is included. Future improvements depend on **larger, well-annotated databases** and the ability to distinguish AF from **mimicking rhythms**.

4.3 Rhythm and Morphology Based AF Detection

Why Rhythm-Based Detection Is Preferred

- Although AF affects both **rhythm** and **atrial wave morphology**, most detectors rely on **rhythm-based** analysis:
  - Rhythm measures (e.g., RR intervals) are **more robust to noise**, especially in wearable and ambulatory recordings [38, 86].
- Rhythm-based detectors:
  - Handle **ectopic beats**, **atrioventricular block**, and patients on **ventricular rate-controlling medication** better.
  - Less prone to **false alarms** in challenging scenarios.

Adding Morphology (P wave, f wave) Improves Specificity

- If **P waves are absent** and/or **f waves are present**, it improves AF detection accuracy.
- Morphological info must be integrated **sensibly**, as shown in the block diagram in Fig. 4.10a (not shown here).
- Important note:
  - **Lead selection** matters:
    - **P waves** are more sensitive to lead position.
    - **f waves** (weaker) are even more dependent on leads **close to the atria**.

Challenges with Morphology-Based Approaches

- Only **a few detectors** use both **rhythm** and **morphology**.
- Their **performance is often inferior** to rhythm-only methods—see Table 4.3 (not shown).
- Main reason: detectors **fail to account for noise**, which varies between leads and conditions.

Design Principle for Morphology Integration

- **Noise level** should guide how much **morphological info** influences detection:
  - If **noise is high**, morphology should **contribute less**.

- If **noise is high**, morphology should **contribute less**,
- If **noise is low**, morphology can be used **more confidently**.
- The system should **reduce reliance** on atrial activity metrics **unless signal quality is high**.

Conclusion

- **Rhythm-based detection remains dominant** due to its **noise resilience**.
- Morphology-based cues (P and f waves) can **reduce false positives**, but:
  - Must be **adaptively weighted** based on **signal quality**,
  - Require **smart integration** to avoid hurting performance.

Fig. 4.10 – Integration of Atrial Wave Morphology and Noise Awareness in AF Detectors

---

(a) Traditional Morphology-Based AF Detector

- **Structure**:
  - ECG → Preprocessing →
    → **RR irregularity**
    → **Atrial information**
    → **Classifier**
- **Limitation**:
  - The **classifier treats all inputs equally**, regardless of **signal noise level**.
  - This can lead to **misclassification** when **atrial signals are unreliable**, especially in noisy recordings.

(b) Noise-Aware AF Detector

- **Structure**:
  - ECG → Preprocessing
    → Branches into:
    - **RR irregularity**
    - **Atrial information**, which includes:
      - **f wave presence**
      - **P wave absence**
    - **Noise level estimation** → All inputs go to a **noise-adaptive classifier**, which:
    - **Reduces reliance** on atrial info as **noise increases**.

- **Advantage**:
  - More **robust to variable recording conditions**, such as:
    - **Ambulatory monitoring**
    - **Wearable ECGs**
  - Ensures **morphological cues** are only used when they're **trustworthy**.

Conclusion:
The (b) model demonstrates a **smart integration** of morphology with **adaptive noise handling**, aligning with current best practices for **reliable AF detection** in real-world noisy environments.

4.3.1 P Wave Detection Information

Use of P Wave Morphology in AF Detection

- **P wave delineation** is well-studied in diagnostic ECG interpretation, where precise measurement of **amplitude and duration** is crucial [91–93].
- A related application is **predicting risk of AF onset** using P wave morphology.

Relaxed Requirements for AF Detection

- In AF detection, **precise P wave boundaries** are **not required**.
- Simply detecting the **absence of P waves** can indicate AF.
- This makes the task less demanding than full P wave delineation [94–96].

Measuring P Wave Absence

- A **straightforward approach**:
  - Analyze **similarity between consecutive PR intervals** using:
    - **Correlation coefficient**
    - **Mean square difference**
  - High similarity → **normal sinus rhythm** (consistent P waves),
  - Low similarity → **AF** (P waves absent or desynchronized from QRS).
- In AF:
  - P waves are replaced by **f waves**, which are not phase-locked with the QRS.
  - So PR intervals vary more → **lower similarity**.
- The **average similarity** across all beats in the detection window can be:

- Compared to a **threshold** to determine AF presence.

Table 4.3 – Performance of Morphology-Inclusive Detectors

Method	Year	Database	Se (%)	Sp (%)
Dash et al. [32]	2009	AFDB$_1$	94.4	95.1
Lian et al. [33]	2011	AFDB	95.8	96.4
Lake and Moorman [34]	2011	AFDB	91	94
Huang et al. [35]	2011	AFDB	96.1	98.1
Shouldice et al. [36]	2012	AFDB	92	96
Lee et al. [37]	2013	AFDB$_1$	98.2	97.7
Zhou et al. [38]	2014	AFDB	96.9	98.3
Asgari et al. [39]	2015	AFDB	97.0	97.1
Petrėnas et al. [40]	2015	AFDB	97.1	98.3
Zhou et al. [41]	2015	AFDB	97.4	98.4
Babuczki et al. [87]	2009	AFDB$_2$	93	98
Carvalho et al. [83]	2015	AFDB$_2$	93.8	96.1
Ladavich & Ghoraani [88]	2015	AFDB$_3$	98.1	91.7
Ródenas et al. [89]	2015	AFDB$_4$	96.5	94.2
Xia et al. [90]	2018	AFDB$_4$	98.3	98.2

- The best-performing detectors often use **rhythm and morphology jointly**.
- **Database differences** (AFDB$_{1-4}$) account for exclusions and dataset balancing decisions (see caption).

Conclusion:
P wave absence can be detected without full delineation, offering a practical way to enhance rhythm-based AF detection. Feature comparisons based on **PR interval similarity** provide a reliable method for detecting the **disruption of atrial activity** that defines AF.

Template-Based P Wave Detection and PQRST Cancellation

Template Matching Approach

- A **PR interval template** is created by averaging many annotated P waves from a **high-quality database** [83, 97, 98].
- In a detection window:
  - **All beats** are compared to the template via **correlation**.
  - A P wave is considered **present** if correlation > fixed threshold.
  - **P wave absence** is defined when the **occurrence ratio** (detected P waves / total beats) falls below a second threshold.

Quantifying P Wave Absence via PQRST Cancellation

- An alternative to detecting P waves directly:
  - Construct a **residual signal** by subtracting the PQRST components from the ECG.
  - This residual contains:
    - **f waves in AF**,
    - Minimal residuals in **normal sinus rhythm**.
- This concept forms the basis of an **echo state network** [100]:
  - Effectively handles:
    - **P wave morphology variability**
    - **Ectopic beats**
  - The network creates an "imaginary" PR interval signal for further analysis.

All-Pair PR Interval Comparison Method

- Builds on [87] by:
  - Evaluating **all pairwise combinations** of PR intervals in the window (not just adjacent ones),
  - Computing **squared errors** between intervals,
  - Averaging all results to form a **global similarity score**.
- The analysis uses:
  - **Fixed PR interval length** = 250 ms,
  - **P wave window** = 250 ms as well.

Conclusion:
Template-based detection and PQRST cancellation offer powerful alternatives for assessing **P wave presence** or absence, enabling **morphology-based AF detection** even in noisy or ectopic conditions. These techniques prioritize **signal similarity** and **residual analysis** over direct waveform detection,

increasing robustness.

Morphology-Only and Wavelet-Based Approaches for AF Detection

---

P Wave–Only Detection Approaches

- Some detectors **exclude rhythm features entirely** and focus solely on **P wave analysis** [88, 89].
- **Motivation**:
  - Rhythm irregularity may be **unreliable** in:
    - Patients with **rate-controlled AF**,
    - Those with a **pacemaker**.
- **Findings**:
  - Table 4.3 shows that such morphology-only detectors still **perform worse** than rhythm-based systems.

Complex Morphology Description Using P Waves

- Example system [88] uses:
  - 6 amplitude-based features (e.g., sampled every 20 ms),
  - 3 shape-based features (variance, skewness, kurtosis),
  - Sampling window is fixed from **QRS fiducial point**.
- **Classifier**:
  - Based on a **Gaussian mixture model**.
  - Trained using **30 minutes of sinus rhythm ECG**,
  - Tests involve computing **Mahalanobis distance** between patient-specific P wave features and candidate waveforms.
- This **novelty detection** setup is well-suited to spotting deviations from the normal patient profile [102, 103].

Wavelet Entropy as Morphology Feature

- Introduced in [89], this method evaluates **entropy across wavelet scales**:
  - Each TQ segment undergoes **wavelet decomposition**:

$$E_i = \frac{\sum_{k=0}^{K_i-1} w_{i,k}^2}{\sum_{l=1}^{J}\sum_{k=0}^{K_l-1} w_{l,k}^2}, \quad i = 1, \ldots, J \tag{4.37}$$

  - E_i represents **relative energy** at each scale.
- Then, **Shannon entropy** of the energy distribution is computed.
- **Results**:
  - **Lower entropy** in intervals with **P waves** (energy more concentrated),
  - **Higher entropy** in **f wave** intervals (energy spread across scales).

Conclusion

- Pure morphology-based detectors, especially those relying on **P wave similarity** or **wavelet entropy**, can be useful alternatives in **rate-controlled or paced** patients.
- However, **rhythm-based** methods still offer superior overall performance in typical AF detection settings.

Challenges of Using P Wave Absence in AF Detection

---

1. PR Interval Variability as a Surrogate

- In AF, the onset of f waves (instead of P waves) leads to **high PR interval variability**.
- This has been considered as a surrogate marker for P wave absence [87].
- However:
  - Determining the PR interval requires detection of **P wave onset and QRS onset**,
  - Since this is **not possible in AF**, the **value of PR interval variability** is questionable.

---

2. Comparison of P Wave Absence Detection Techniques

- Techniques vary greatly:
  - From **simple similarity measures** (e.g., correlation, mean square error),
  - To **advanced statistical modeling**.

Similarity-Based Approaches:

- Don't favor specific P wave morphology (polarity or shape),
- Work well across **varied rhythms**.

Template-Based Approaches:

- Can struggle when P wave **morphology differs from the template**,
- The correlation drops when P wave and template are "orthogonal."

Statistical Modeling:

- More flexible and accurate,
- But requires **extensive sinus rhythm training data** per patient,
- Often **infeasible** in practical settings.

3. Sensitivity to Noise Levels

- All P wave absence detection becomes **increasingly unreliable** as **noise increases**,
- Each technique has a **noise "breakdown" threshold**, which varies by design.

Example:

- **Template-based detection** is more robust to noise than similarity measures.
- Using **multiple ECG leads**:
  - Enhances P wave visibility,
  - Improves performance: e.g., **sensitivity rose from 91.7% to 94.6%** in [88].

Conclusion:
P wave absence is a valuable cue for AF detection, but it's **highly sensitive to noise** and **P wave morphology variations**. Simpler similarity-based methods offer robustness across morphologies, while template-based and statistical models can improve accuracy if clean, long-duration training data are available. Using **multi-lead analysis** provides a practical way to improve reliability.

4.3.2 f Wave Detection Information

Challenges in Detecting f Waves

- f waves have **low amplitude** and can be easily **masked by noise**.
- The **TQ interval** (used for f wave analysis) becomes **shorter at higher heart rates**, making detection more difficult.
- At very high heart rates:
  - f waves may become **indistinguishable** from baseline fluctuations or muscle noise,
  - Or be entirely **buried**, especially in **wearable ECGs**.

Approach Described in [104]

- **Method**:
  - Count the **number of f waves** in the TQ interval.
  - A wave is identified if:
    - A **signal fluctuation** exceeds a **fixed threshold**,
    - The **width** between two zero-crossings exceeds a second threshold.
- **Noise Mitigation**:
  - The method includes:
    - **Adaptive thresholding** based on **TQ duration** and **peak amplitude**.
    - Preprocessing steps to remove **baseline wander** and **muscle noise**.
- **Limitation**:
  - Even after filtering, f wave analysis based on **level crossings** remains **noise-sensitive**.
  - Muscle noise can **mimic the frequency content** of f waves.
  - **High heart rates** result in **short TQ intervals**, increasing the chance of **false negatives** (zero f waves counted).

Conclusion

- While f wave presence can be a useful signal for AF detection, it is:
  - Highly **susceptible to noise**,
  - Sensitive to **heart rate increases**,
  - Difficult to extract without **complex processing**,
- Therefore, f wave detection may be **impractical** for **low-power or wearable devices**, unless advanced signal extraction methods are applied.

Spectral Characterization of f Waves

---

1. Dominant Frequency Analysis

- f wave presence can be inferred by analyzing **spectral content** of the extracted atrial signal.
- This approach assumes **successful f wave extraction** across the entire detection window (not just TQ interval) [83, 100].
- The **dominant atrial frequency (DAF)**, typically near **6 Hz**, is the most prominent feature in the spectrum (see Fig. 4.11a, not shown).
- Additional info may come from:
  - **Second and third harmonics**,
  - More commonly seen in **paroxysmal AF** than in **permanent AF**.

---

2. Normalized Spectral Concentration (F_{SC})

- Used to **quantify the presence of concentrated spectral energy** around DAF:

$$F_{\text{SC}} = \int_{\Omega_a} P'_j(\omega)\, d\omega \tag{4.38}$$

- Where:
  - $P'_j(\omega)$ = **normalized power spectrum** of the extracted f wave signal $\hat{d}(n)$:

$$P'_j(\omega) = \frac{1}{\sigma^2_{\hat{d}}} P_j(\omega) \tag{4.39}$$

- $\sigma^2_{\hat{d}}$: variance of the extracted signal $\hat{d}(n)$,
- Ω_a: frequency range centered around DAF (e.g., 4–12 Hz),

- If **f waves are present**, $F_{\mathrm{SC}} \approx 1$; otherwise, it is lower (e.g., during sinus rhythm).
- Power spectrum estimation can use:
  - **Welch's method** (nonparametric),
  - **Burg's method** (parametric) [107].

3. Spectral Entropy (Fₛₑ)

- Measures **distribution of spectral power** to assess **regularity/complexity**:

$$F_{\mathrm{SE}} = -\int_{\Omega_a} P_j^{\prime}(\omega) \ln\left(P_j^{\prime}(\omega)\right) d\omega \tag{4.40}$$

- Lower entropy → **sharp peak** (f wave present),
- Higher entropy → **flat spectrum** (f wave absent or buried in noise) [83].

Conclusion:
Spectral concentration and spectral entropy are reliable tools for identifying **f wave presence**, especially when a clear DAF can be extracted. They provide a **quantitative spectral signature** of AF, particularly effective when noise levels are manageable and f wave extraction is successful.

Fig. 4.11 & KL Divergence for f Wave Analysis

Fig. 4.11: Spectral Comparison of f Wave vs. Sinus Rhythm

- **(a)**: Power spectrum of an **extracted f wave signal**:
  - Shows a **prominent peak at ~6 Hz** (the dominant atrial frequency),
  - Spectral content is **concentrated**, supporting **AF presence**.
- **(b)**: Spectrum of a **QRST-cancelled sinus rhythm signal**:
  - Contains **multiple small peaks**, none dominant,
  - Reflects a **broader, noisier** spectral distribution typical of sinus rhythm.

> Vertical dashed lines indicate the two **largest peaks** in each case.

Kullback–Leibler (KL) Divergence

- The **KL divergence** quantifies the **difference between**:
  - The observed spectrum $P_j^{\prime}(\omega)$ and
  - A **template spectrum** $P_j^{r}(\omega)$ from known f wave data [83].

$$F_{\mathrm{KL}} = \int_{\Omega_a} P_j^{\prime}(\omega) \ln\left(\frac{P_j^{\prime}(\omega)}{P_j^{r}(\omega)}\right) d\omega \tag{4.41}$$

- **Interpretation**:
  - $F_{\mathrm{KL}} \approx 0$: Observed spectrum matches the template → **f waves likely present**,
  - High F_{KL}: Spectral distribution deviates → **f waves likely absent**.

Limitations of KL Divergence

- The **template spectrum** $P_j^{r}(\omega)$ must:
  - Be **representative across patients**, despite:
    - Variability in **DAF** (dominant atrial frequency),
    - Differences in **f wave morphology**.
- In [83], the **template** was built from AFDB data and used with **spectral entropy** for detection.
- However:
  - DAF may vary between **4 and 12 Hz**,
  - A mismatch in DAF or morphology may limit effectiveness of KL divergence.

Conclusion:
KL divergence provides a powerful, relative measure of spectral similarity to known f wave profiles, but its **success depends heavily on the generalizability of the spectral template**. When paired with **spectral entropy** and robust **f wave extraction**, it adds an extra layer of confirmation for AF detection.

Spectral Structure Tests for f Wave Detection

Heuristic Spectral Tests (proposed in [108]):

These tests determine if the extracted power spectrum $P_j^{\prime}(\omega)$ resembles a typical **AF spectrum**, based on 5 heuristic checks:

1. **Signal-to-Noise Ratio (SNR)**:
   - "Signal": average of the **two largest harmonics**.
   - "Noise": average around midpoint between the harmonics.
2. **Second Harmonic Deviation**:
   - Check if the **second peak** is where expected based on first harmonic

- Check if the **second peak** is where expected based on first harmonic,
- Helps exclude "ringing" spectra (e.g., due to slow P waves).

3. **Magnitude Ratio**:
   - Ratio of **second to first harmonic** magnitude,
   - Validates presence of **multiple harmonic structure**.
4. **Spectral Sharpness**:
   - Compare the sliding window spectrum to an **exponentially averaged background**,
   - Helps suppress **muscle artifacts** or **poor f wave extraction**.
5. **Total Residual Energy**:
   - Check for **excess power** outside harmonic bands,
   - High residuals → likely **non-AF signal**.

Interpretation of Fig. 4.11a vs. 4.11b

- **Fig. 4.11a**: Passes all tests → classified as **AF spectrum**.
- **Fig. 4.11b**: Fails at least one test → **not AF**.

Fuzzy Logic Combination of P and f Wave Info

- Introduced in [100], tested in Fig. 4.12:
  - Uses both **P wave absence** and **f wave presence**,
  - In example with APBs and respiratory sinus arrhythmia:
    - Detector based on **fuzzy logic** correctly rejects **non-AF rhythms**,
    - Detector using I_{CSampEn} (coefficient of sample entropy) falsely detects AF.

Conclusion:
The proposed spectral structure tests are **non-model-based**, computationally light, and effective for screening AF-like spectra. When combined with **morphological features** (e.g., in fuzzy logic detectors), they greatly enhance AF specificity—especially in challenging cases like APBs or respiratory sinus rhythm.

Fig. 4.12 – Avoiding False AF Detections with Morphology-Based Detection

---

Overview of Figure 4.12

- Compares two detection methods on **non-AF rhythms** that commonly trigger false alarms:
  - **(a)** Frequent Atrial Premature Beats (APBs),
  - **(b)** Respiratory Sinus Arrhythmia.

For each case:

- **Top trace**: ECG signal.
  - In (a), * indicates each APB.
- **Middle trace**: Rhythm-based AF detection output Q_R (from [34]).
- **Bottom trace**: Rhythm + morphology-based AF detection output Q (from [100]).

AF is detected whenever the decision function (thicker line) exceeds the threshold.

Key Observations

- **Rhythm-only detection (Q_R)**:
  - Triggers **false positives** in both (a) and (b),
  - Misclassifies **non-AF rhythms** as AF due to rhythm irregularity alone.
- **Rhythm + Morphology detection (Q)**:
  - Remains **below threshold** in both examples,
  - Correctly identifies that **AF is not present**, despite irregularities.

Conclusion

This figure highlights the **importance of incorporating atrial morphology** (e.g., P wave presence, f wave absence) in AF detectors. While rhythm-based methods alone can misinterpret **benign irregular rhythms**, the combination with morphology **substantially reduces false alarms**, especially in complex or noisy clinical scenarios.

4.3.3 Noise Level Estimation

Problem Addressed

- Many AF detectors **ignore noise levels** entirely.
- Most ECG signals are processed **uniformly**, regardless of noise.
- This can lead to **discarding morphology information** (e.g., f waves, P waves) under noisy conditions (as shown in Fig. 4.10).
- The challenge lies in:
  - Designing a **noise estimator** that doesn't confuse noise with cardiac activity,
  - **Integrating noise estimation into decision logic**.

Noise-Adaptive Detection ([100])

- An innovative detector adjusts to noise using the extracted **f wave signal** $\hat{d}(n)$.
- It estimates noise based on **spectral entropy** from $\hat{d}(n)$ as follows:

$$\hat{N}_{\mathrm{RMS}} = R_d \cdot \frac{\int_{\Omega_n} P_{\hat{d}}(\omega) \log_2 P_{\hat{d}}(\omega) d\omega}{\int_{\Omega_a} P_{\hat{d}}(\omega) \log_2 P_{\hat{d}}(\omega) d\omega} \tag{4.42}$$

Details of the Formula

- R_d: Root mean square (RMS) of the extracted atrial signal $\hat{d}(n)$,
- $P_{\hat{d}}(\omega)$: Power spectrum of $\hat{d}(n)$,
- Ω_n: Frequency range **dominated by noise** (e.g., outside f wave range),
- Ω_a: Frequency range **dominated by f waves** (e.g., around 4–12 Hz),
- The **ratio of entropies** compares:
  - Broad, high-entropy **noise-like spectra** in Ω_n,
  - Focused, low-entropy **f wave spectra** in Ω_a.

> **Interpretation:**

- Low $\hat{N}_{\mathrm{RMS}}$ → **Low noise** (f waves dominate),
- High $\hat{N}_{\mathrm{RMS}}$ → **High noise** (muscle/motion artifacts dominate).

Purpose

- Enables AF detectors to **discard unreliable morphology features** when noise is high,
- While **preserving sensitivity** during clean recordings.

Conclusion:
This approach introduces a dynamic, data-driven method to estimate noise that **doesn't rely on visual inspection or fixed thresholds**. It allows **morphology-based detectors** to adapt their behavior depending on signal quality—ensuring **reliable AF detection even in wearable or ambulatory settings**.

Fig. 4.13 – Noise Estimation Using Spectral Entropy

Description of Subfigures

- **(a)** ECG segment with:
  - Two **AF episodes**,
  - Middle section (15–25 s): contaminated by **myoelectric noise**,
  - Two **atrial premature beats (APBs)** marked by "*" before second AF episode.
- **(b)** Extracted **f wave signal** $\hat{d}(n)$ from echo state network.
- **(c)** Noise level estimate $\hat{N}$ computed using Eq. (4.42) with a **5-beat sliding window**.

Key Observations

- The **first 15 seconds** are noise-free.
- **Myoelectric noise** significantly increases the value of $\hat{N}$ starting around 15 s.
- Second AF episode remains **undistorted by noise**, and thus noise estimator **returns to low levels**.

Interpretation

- The estimator $\hat{N}_{\mathrm{RMS}}$ clearly **tracks the onset and offset** of the noise burst,
- It **remains low** during clean AF segments, suggesting good **specificity**,
- Shows that noise levels can be estimated **independently** of cardiac activity using the power spectrum and entropy of $\hat{d}(n)$.

Conclusion

This figure validates the **noise estimator** from Eq. (4.42) as an effective tool for detecting **myoelectric noise contamination**. Such an estimator can inform AF detectors to **adjust sensitivity** or **discard unreliable morphology features**, especially in wearable/ambulatory ECG monitoring.

4.3.4 Ectopic Beat Handling

Challenge

- Ectopic beats (especially APBs and VPBs) pose a **major issue for AF detection**, leading to **false alarms**.
- Particularly problematic for **rhythm-based detectors**, which may confuse irregularity from ectopic beats with AF.

Indirect Handling through Morphology

- Detectors using:
  - **P wave absence** ([87], [88]) or
  - **P wave absence + f wave presence** ([100])

  can **reduce false detections** of APBs.
- Key condition: APB must be **preceded by a detectable P wave**, even if its morphology differs from normal sinus beats.
- If APB's P wave is **hidden in a T wave**, the risk of false AF detection **increases**.

VPBs and Risk of Misclassification

- VPBs are **not preceded by P waves**, increasing their risk of **false AF detection**, particularly in **rhythm-only detectors**.
- Morphology-aware detectors (checking for **f waves and P wave absence**) perform **better in ectopic conditions**.

Built-in Morphology Classifiers

- For automated ECG systems (holter/implantable):
  - **Morphology classification** is often available,
  - Can be leveraged to **flag or exclude segments** with VPBs before running AF detection,
  - May also assist by **augmenting the feature vector** used for classification ([83]).

Conclusion

Incorporating beat morphology into AF detection systems—especially those already designed for long-term ECG monitoring—offers a **reliable strategy** for managing **ectopic rhythms**. Rather than relying solely on rhythm irregularity, using **P wave presence/absence and f wave analysis** significantly **reduces false positives** from APBs and VPBs.

4.3.5 Classification

Rhythm and Morphology-Based AF Detection

- Principles from rhythm-only classification (see Sect. 4.2.6) also apply here.
- Now, **morphology-derived features** like **P wave absence** and **f wave presence** are included.
- **Noise level** should ideally be included in the feature vector to help **assess reliability**, but:
  - Many classifiers still do **not explicitly use noise estimates**,
  - Instead, they are trained on datasets **with varied noise levels** to generalize.

Early Detector Examples

- One of the first combined rhythm + morphology AF detectors was from [87]:
  - Feature vector:
    - One rhythm parameter (Markov-based),
    - Two P wave-related parameters (P wave similarity, PR interval variability).
  - Classifier: regression decision tree using **threshold tests** only.

Multivariate Mixture Model Approach

- Described in [88] using **9-dimensional P wave amplitude features**:
  - Each class (AF vs. non-AF) modeled using **sum of Gaussians** (PDF),
  - Parameters estimated via **expectation-maximization**,
  - Requires a **patient-specific training phase**.
- Classification:
  - For each beat, compute **likelihood** of P wave features,
  - Decision based on **combined likelihood** from all beats in the detection window.

Additional Methods

- Other approaches:
  - **SVM-based classification** using P wave features ([83], [121]),
  - **ANNs** and **regression tree classifiers** used for P wave absence + f wave presence,
  - **Decision fusion strategies** considered to merge multiple feature outcomes.

Summary

Classification methods for AF detection with rhythm and morphology features range from:

- **Simple threshold-based rules** (e.g. regression trees),
- To **complex probabilistic models** (e.g. Gaussian mixtures),
- And **machine learning models** (e.g. SVMs, ANNs).

Patient-specific models improve accuracy but demand more computation and training data.
Inclusion of **noise robustness** remains a key future direction.

4.3.5 Classification (continued)

Rhythm vs. Rhythm and Morphology-Based Detection

- **Performance comparison (Table 4.3):**
  - Detectors using **only rhythm** can outperform detectors using **rhythm + morphology**.
  - Example: detector in [31] (rhythm-only) outperforms [83] (which includes P and f wave info).
  - Reason: **decision boundaries not adapted to varying noise levels**.

Impact of Noise

- Studies ([83], [87]–[89]) all emphasize noise as a major issue degrading performance.
- **None** of the detectors were explicitly designed to **account for noise**.
- Most classifier structures (ANNs, regression trees, Gaussian mixtures) **do not generalize well** from low to high noise conditions.

First Noise-Aware AF Detector

- Introduced in [100]:
  - Four parameters used: RR irregularity, P wave absence, f wave presence, and **noise level** (from Eq. 4.42).
  - Classifier: **Mamdani-type fuzzy logic inference system**.
    - Each parameter is fuzzified into levels (e.g., "low", "high").
    - 16 **if-then rules** combine the fuzzified values.
    - The fuzzy inference yields a decision value $O \in [0, 1]$,
      - **Closer to 1** = AF likely,
      - **Closer to 0** = non-AF.
- This fuzzy logic system is visualized in **Fig. 4.10b**, reflecting:
  - **Noise-dependent behavior**,
  - **Balanced weighting** of all four parameters.

Summary

- Many advanced AF detectors **do not surpass** rhythm-only ones, mainly due to:
  - Lack of adaptation to noise,
  - Overfitting to clean training data,
  - Inflexible decision boundaries.
- **Fuzzy logic-based detectors** show promise by **incorporating noise estimates** and interpreting features **in a non-binary fashion**.

This concludes the section on classification within AF detection using both rhythm and atrial wave morphology.

Final Remarks on Noise-Aware Classification

Advantages of the Fuzzy Logic Classifier

- **No training phase required**.
- Rules and membership functions are based on **prior knowledge of AF**.
- However, its performance has **not been validated on AFDB**, as f wave extraction demands high-quality signals (which many AFDB recordings lack).

Alternative Approach: Beat Exclusion

- Suggested in [89]: **Exclude beats** if their estimated noise level exceeds a fixed threshold.
- Thresholds are tuned to match **manual annotations of noisy segments**.
- In noisy ECG segments:
  - **Detector operation is suspended** when P wave absence can't be determined.
  - This contrasts with the fuzzy logic system in [100], which **continues to operate**, relying more on **rhythm irregularity** under high noise conditions.

Figure 4.14

This figure extends Fig. 4.13 and illustrates **noise-aware fuzzy logic AF detection** using the system in [100]. It shows:

1. **Top trace**: ECG signal with marked AF and noise segments.
2. **Second trace**: Extracted f wave.
3. **Bottom five traces**: Trends of the input parameters:
   - $\hat{N}$: Estimated noise level
   - R: Rhythm irregularity
   - F: f wave presence
   - P: P wave absence
   - O: Final output decision (AF likelihood)

- O: Final output decision (AF likelihood)

This shows how **noise level affects the contribution** of morphology-based parameters (P, F), with the system relying more on R under high noise.

Comment on the Most Recent Rhythm and Morphology-Based Detector (Table 4.3)

- **Best performance** among detectors in Table 4.3 is achieved by the **most recent model**, which slightly outperforms all others.
- This detector is based on a **deep convolutional neural network (CNN)**.
  - Input: either **short-term Fourier transform (STFT)** or **stationary wavelet transform** of 5-second ECG segments.
  - The input includes **both atrial and ventricular activity**.
- Unlike previous methods, **the design is not physiology-based**.
  - No explicit modeling of rhythm irregularity, P wave absence, or f wave presence.
  - Instead, general **signal properties** and possibly **visual features** (e.g., color vs. greyscale) of the time-frequency representations are emphasized.

Caveats

- Despite promising performance, the results **must be questioned** due to:
  - Use of **only a subset of the AFDB**.
  - Application of **tenfold cross-validation**, which may limit generalizability.
  - These concerns are expanded on in **Section 4.6**.

4.5 Performance Measures

- Detection performance for AF (atrial fibrillation) is quantified by comparing detected beats to annotated beats. Four key counts used:
  - N_{TP}: Beats correctly identified as AF (true positive)
  - N_{TN}: Beats correctly identified as non-AF (true negative)
  - N_{FP}: Beats incorrectly identified as AF (false positive)
  - N_{FN}: Beats incorrectly identified as non-AF (false negative)
- Common performance measures include:

  Sensitivity:

  $$\text{Sensitivity} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (4.43)$$

  Specificity:

  $$\text{Specificity} = \frac{N_{TN}}{N_{FP} + N_{TN}} \quad (4.44)$$

- Performance is often visualized via ROC (Receiver Operating Characteristic) curves, plotting sensitivity vs. (1-specificity) at varying detection thresholds (see references [34, 37]).
- ROC performance is summarized using the Area Under the Curve (AUC), where:
  - AUC = 1.0 → perfect performance
  - AUC = 0.5 → random performance
- In AF detection, parameters can be optimized by maximizing AUC (see references [39, 41, 51, 88]).
- Additional measures include:

  Positive Predictive Value (PPV):

  $$\text{Positive predictive value} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (4.45)$$

  Detection Accuracy:

  $$\text{Detection accuracy} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FN} + N_{FP} + N_{TN}} \quad (4.46)$$

Sensitivity and Specificity: Episode-Based Comparison

- Traditional sensitivity and specificity, based on beat-to-beat comparison, don't reflect the episodic nature of paroxysmal AF accurately.
- Example given:
  Two AF episodes—one hour-long and another brief 10-beat duration.
  A detector might detect the long episode but miss the short one, thus sensitivity remains high despite missing a full episode.
  Conversely, a short false-positive episode has negligible effect on specificity.
- This illustrates the limitation of beat-to-beat performance metrics in representing clinical relevance, as brief episodes can be overlooked, thus losing significant clinical information.
- The ROC curve might misleadingly indicate near-perfect performance, as depicted in Fig. 4.15.

Alternative Approach: Episode-to-Episode Comparison

- Suggestion: Replace beat-to-beat comparison with episode-based comparison.
- Challenges/questions arising:
  - How should "true negative" be defined in the absence of an annotated AF episode?
  - Should minimum duration criteria define a detected episode as correct?
  - Treatment of episodes labeled AF beats only partially?
- Inspired by performance measures in transient ischemia detection for long-term ECG (reference [126]), adapted for AF detection ([36]):
  - Episodes of non-AF beats have limited clinical meaning, making true negatives (N_{TN}) undefined.
- Thus, counts are redefined as:
  - N_{TP}: AF episodes correctly detected (true positive)

- N_{FP}: Non-AF episodes falsely detected as AF (false positive).
- N_{FN}: AF episodes falsely detected as non-AF (false negative).

Episode Detection Criteria and Performance Illustration (Fig. 4.15)

- An AF episode is considered **correctly detected** if it overlaps at least **50%** with the annotated episode; otherwise, it is labeled non-AF [33].

Figure 4.15 Description:

- **(a)** Shows an RR interval time series $x(n)$.
- **(b)** Annotated episodes: single AF episode clearly marked; rest is non-AF.
- **(c)** Detector output based on coefficient of sample entropy (using a 12-beat window), detecting numerous false positives due to ectopic beats.
- **(d)** Corresponding ROC curve indicates near-perfect detection despite many brief false-positive episodes, illustrating a limitation of beat-based ROC metrics.

Performance Impact of Window Length:

- For a **100-beat window**, sensitivity is high (**0.92**) using beat-based comparison ([36], Table 4.1).
- When using an **episode-based sensitivity** measure, sensitivity drops significantly (**to 0.71**), demonstrating reduced detection of short episodes.

This emphasizes the importance of choosing appropriate detection criteria (episode-based rather than beat-based) to avoid overstating detector performance.

Episode-Based vs. Beat-Based Performance Measures and Detection Delay

- **Episode-based measures** are not yet widely adopted in AF detection literature, despite providing complementary information to beat-based measures.
- The popularity of **beat-based measures** arises from their computational simplicity and well-established use in ECG applications.
- Neither beat-based nor episode-based measures provide insights on detecting episodes of varying durations effectively.

Detection Delay as an Additional Performance Measure:

- **Delay** between annotated episode onset and detector-produced onset has gained attention ([32, 35, 39, 89]).
- From an algorithmic perspective, defining and assessing this delay is crucial for comparing performance with annotated episode boundaries.
- Clinically, however, a short detection delay is typically less important than accuracy-based measures, as few ECG applications require immediate response upon episode initiation.

4.6.2 Training and Evaluation

- Approaches to classifier **training and performance evaluation** vary significantly, as evidenced by studies summarized in Table 4.3.
- Different studies may:
  - Use a **proprietary database or LTADFB** for training and evaluate performance using AFDB [38, 40, 41, 87].
  - Avoid using the **same patients** for training and evaluation.
  - Omit details about training dataset sources entirely [33, 37].
  - Train and evaluate using AFDB or other publicly available databases.

Usage of AFDB:

- AFDB commonly serves to:
  - Determine **optimal detection thresholds** ([32, 34]).
  - Select optimal **feature sets** for classifiers ([36]), followed by evaluation on other databases.
- Training on AFDB introduces a positive **bias**, yet some studies include these biased results in performance comparisons ([38, 88, 89, 128]), despite figures not being fully representative.
- Exception: Reference [35] uses AFDB explicitly for both training and evaluation, making its results clearly comparable.

Methods to Reduce Bias with AFDB:

- Partition AFDB into separate subsets for training and evaluation to avoid bias:
  - **Random selection** of non-AF/AF segments.
  - **Stratified twofold cross-validation** ([39]).
  - **Tenfold cross-validation** ([90]).
- Example: A subset of AFDB was used for training ([83]), and evaluation was conducted on the full dataset afterward, and vice versa.
- While cross-validation on AFDB can assess variability, it remains debatable if these figures generalize to separate datasets due to the limited size of AFDB ([129]).

Patient-based Training and Evaluation Considerations

- Training and evaluation methods mentioned earlier are usually population-based; however, a

patient-based approach can also be employed ([88]).

- In patient-based training:
  - Training is based on the initial portion of an ECG recording, and evaluation on the remaining part.
  - Beats containing **irregularities** are manually excluded from training, inadvertently introducing positive bias.
  - Practical limitations exist, as manual review and good-quality signals aren't always feasible in routine clinical settings.
- Comparing detection performance is significantly challenged by **positive bias** when the same patient data is involved in both training and evaluation.
- Ideal evaluations involve:
  - Independent datasets for training and evaluation.
  - Different patient populations for training vs. testing.
- Due to positive bias, comparisons of detection performance metrics (sensitivity and specificity) must be interpreted cautiously, as highlighted in Table 4.3.

Special Note on Adaptive Filtering-Based AF Detection:

- AF detectors using **adaptive filtering** for f-wave extraction ([100]) can't reliably train and evaluate using AFDB, because:
  - AFDB does not contain leads suitable as a pure reference, lacking leads with negligible atrial activity.
- Solutions include:
  - Training on a proprietary multi-lead database.
  - Evaluating using simulated multi-lead signals ([100]).

4.6.3 Simulated ECG Signals

- Typically, AF detector performance is evaluated using real ECG signals annotated for AF episodes. However, simulated ECG signals are used infrequently.
- In contrast, for **f-wave extraction**, simulated signals are preferred because manual annotations are irrelevant to the extraction process.
- Simulated ECG signals are advantageous for studying clinically and technically significant properties, such as:
  - **Atrial ectopy**
  - **Episode duration**
  - **Noise levels**

Detection Accuracy and Simulated ECG Studies ([100])

- Detection accuracy was studied using simulated signals with varying noise levels, both **with and without Atrial Premature Beats (APBs)**.
- Findings highlighted:
  - Importance of properly handling APBs.
  - Detectors using **both rhythm and morphology** yield much higher accuracy compared to rhythm-only detection in presence of APBs.
  - **Lower noise levels** facilitate accurate morphology-based detection (due to reliable estimation of f-waves).
- Similar accuracy differences are noted in detecting episodes of varying lengths:
  - **Rhythm and morphology-based detection** is superior for brief episodes (5, 10, 20, 30 beats).

Simulated ECG Signals for SNR Evaluation

- Simulated signals help determine the minimum SNR (Signal-to-Noise Ratio) at which AF detection remains reliable.
- Example study ([114]):
  - Muscle noise was simulated and added to real ECGs from LTADFB at different SNRs.
  - Results showed minimal reduction in detection accuracy when SNR was expressed in terms of decibels (dB).
- Conclusion:
  - Noise level evaluations using simulated signals are essential, particularly for AF detectors analyzing both **rhythm and morphology**.

4.6.4 Brief AF Episodes

- Clinical importance exists for detecting brief **Paroxysmal AF (PAF)** episodes due to associated stroke risks, yet minimal attention is given to short episodes in detector evaluations.
- Typically, AFDB includes mostly **longer AF episodes**, with brief episodes having minimal effect on beat-based performance measures.
- However, some studies highlight missed brief episodes:
  - In [32], **30 out of 254 episodes** were missed, all shorter than 75 beats.
  - In [35], **32 out of 299 episodes** were missed, primarily due to their brief durations (4 to 62 beats).

Detection Window Length and Brief Episodes

- Performance indirectly evaluated via **detection window length** analysis:
  - Window length defines minimum detectable AF duration.
  - Shorter windows facilitate detection of brief episodes but increase false detections.
  - Longer windows improve rhythm estimation but miss shorter episodes (illustrated in **Fig. 4.16**).

Figure 4.16 Explanation:

- **(a)** Shows RR interval series from AFDB (record 4043), featuring a short AF episode (20 beats) followed by a longer one.
  - Detection window (sliding box) is too wide, missing the initial brief episode.
- **(b)** Depicts manual annotation vs. detector output, highlighting detection delay and missed brief episode.

Trends in Detection Window Length:

- Historically, detection window length has significantly decreased:
  - From **180 seconds** (1992) [130] to just **8 beats** (2015) [40].
- Further reduced lengths (down to **5 beats**) for rhythm-and-morphology-based detectors have been studied ([100]).

Recommended Detection Window:

- Currently recommended window length:
  - **8 beats** for rhythm-based detectors.
  - **5 beats** for rhythm-and-morphology-based detectors.

Performance Degradation with Short Detection Windows

- Shorter detection windows lead to **reduced performance** in AF detectors based on time-varying coherence ([37], Sect. 4.2.3):
  - **128-beat window**: Sensitivity **98.2%**, Specificity **97.7%** (Table 4.1).
  - **32-beat window**: Sensitivity drops to **96.7%**, Specificity to **96.1%**.
  - Despite reductions, shorter windows may still provide more accurate AF burden estimations.
- For simpler detectors using **Poincaré plot population distribution** ([33]):
  - **128-beat window**: Sensitivity **95.9%**, Specificity **95.4%**.
  - **32-beat window**: Sensitivity **94.4%**, Specificity **92.6%**.

Direct Evaluation with Simulated ECG Signals ([40]):

- Simulated ECGs allow controlled analysis of detector accuracy vs. episode length (**median duration** T_E).
- Results illustrated in **Fig. 4.17**:
  - Detection accuracy decreases with shorter episode durations.

Figure 4.17 Explanation:

- Compares **rhythm-only** vs. **rhythm-and-morphology-based** detectors.
- Two types of signals analyzed:
  - (a) **Synthetic components**
  - (b) **Real components**
- Noise level fixed at **20 µV RMS**.
- Rhythm-and-morphology consistently outperforms rhythm-only, particularly with shorter episode lengths.

Observations from Fig. 4.17:

- Detection accuracy difference between rhythm-only and rhythm-and-morphology-based detectors **increases significantly** for shorter episodes.
- **Real ECG signals** cause larger performance drops in rhythm-only detectors compared to synthetic signals, due to the presence of real pathological rhythms.

4.7 Additional Detector Information

- **Prediction of AF Onset and Offset:**
  - Certain ECG properties explored for predicting AF onset/end:
    - More than 90% of AF episodes are triggered by APBs (atrial premature beats) ([131–135]).
    - Simple test: Detect increase in APBs not followed by regular RR intervals.
    - Additional tests include analyzing runs of **atrial bigeminy/trigeminy** and short runs of **paroxysmal atrial tachycardia** ([136]).
- **Heart Rate Variability (HRV) Analysis:**
  - Paroxysmal AF onset can be preceded by changes in HRV.
  - Significant reduction in low-to-high frequency ratio of HRV ([137–139]) indicates AF onset, not detectable after spontaneous recovery ([140]).
  - HRV changes described by entropy measures suggest AF preceded by RR interval complexity ([141], [142]).
  - AF onset prediction improved using **spectral, bispectral, and nonlinear features**, plus machine learning classification of preceding HRV patterns ([143]).
  - Typically, 30-minute segments analyzed to evaluate APB- and HRV-based prediction.
- **P Wave Morphology Changes:**
  - Abnormal interatrial conduction changes P wave morphology prior to AF ([144–145]).
  - Duration of minimum P wave may predict recurrent AF ([146–148]).
  - Shortening of minimum P wave duration is also predictive ([147, 149]).
  - P wave morphology alone may predict AF onset ([96, 148]).
  - However, P wave properties change over longer durations (weeks-months) compared to short-term APB- and HRV-based predictors (minutes), reducing immediate predictive value

for AF detectors.

- **AF Episode Termination Prediction:**
  - Detection of AF termination (return to sinus rhythm) is complex, as it depends on cancellation of ventricular activity.
  - **Dominant atrial frequency (DAF)** slows just before termination ([150, 151]).
  - Studies using **wavelet entropy** (reflecting time-frequency unpredictability) suggest potential for termination prediction ([152]).
- **Role of Physical Activity:**
  - Physical activity information likely improves AF detection.
  - ECG devices increasingly include **accelerometers** ([153]).
  - Although combined ECG-accelerometer data improves accuracy, preliminary results indicate accuracy reduction when excluding accelerometer data ([153]).
  - Accelerometers alone can detect AF episodes, demonstrated by studies with chest-mounted ([154]) and bed mattress sensors ([155]), without ECG signals.