

Research Article

On Consistency of Cross-Approximate Entropy in Cardiovascular and Artificial Environments

Tamara Skoric,¹ Olivera Sarenac,² Branislav Milovanovic,³
Nina Japundzic-Zigon,² and Dragana Bajic¹

¹Faculty of Technical Sciences-DEET, Center of Excellence CEVAS, University of Novi Sad, Trg Dositeja Obradovica 6, 21000 Novi Sad, Serbia

²School of Medicine, Institute of Pharmacology, Clinical Pharmacology and Toxicology, University of Belgrade, Dr. Subotica 1, 11129 Belgrade, Serbia

³Department of Cardiology, Neurocardiology Laboratory, University Hospital Medical Center "Bezanijska Kosa", Bezanijska Kosa bb, 11080 Belgrade, Serbia

Correspondence should be addressed to Dragana Bajic; dragana.bajic@gmail.com

Received 26 February 2017; Revised 8 June 2017; Accepted 7 August 2017; Published 17 September 2017

Academic Editor: Zoran Obradović

Copyright © 2017 Tamara Skoric et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cross-approximate entropy (*XApEn*) quantifies the mutual orderliness of simultaneously recorded time series. Despite being derived from the firmly established solitary entropies, it has never reached their reputation and deployment. The aim of this study is to identify the problems that preclude wider *XApEn* implementation and to develop a set of solutions. Exact expressions for *XApEn* and its constitutive parts are derived and compared to values estimated from artificial data. This comparison revealed vast regions within the parameter space that do not guarantee reliable probability estimation, making *XApEn* estimates inconsistent. A simple correction to one of the *XApEn* procedural steps is proposed. Three sets of formulae for joint parameter selection are derived. The first one is intended to maximize threshold profile. The remaining ones minimize *XApEn* instability according to the strong and weak criteria. The derived expressions are verified using cardiovascular signals recorded from rats (long signals) and healthy volunteers (short clinical signals), proposing a change of traditional parameter guidelines.

1. Introduction

Approximate entropy (*ApEn*) is among the most exploited nonlinear techniques to quantify the complexity and unpredictability of time series [1–3], with a citation rate highlighted in [4]. *ApEn* is approved as a supporting tool in preclinical and clinical studies, with the most prominent applications in cardiovascular studies.

In spite of the vast and firmly established implementation, the activities on *ApEn* improvement have never ceased. Research efforts are oriented towards modifications, including sample entropy (*SampEn*) [5], corrected conditional entropy (*CAPEn*) [6], *K*-nearest neighborhood entropy (*KNNCE*) [7], fuzzy entropy (*FuzzyEn*) [8], multi-scale entropy (*MultiScaleEn*) [9], distance entropy (*DistEn*) [10], binarized entropy (*BinEn*) [11], *ApEn* based on wave mode [12], and speeding-up algorithms [13, 14], and towards

the consistency studies, including threshold selection [15–20], time lag [21], and consistency in general [4].

One of the very first *ApEn* modifications is *Cross-ApEn* (*XApEn*) [22] that estimates mutual predictability of two simultaneously recorded time series. In view of *ApEn* as a complexity measure of solitary signal, *XApEn* should have become an equivalent tool for paired physiological processes. However, the reputation and deployment of its predecessor have never been reached. Except for a recent improvement [23], *XApEn* applications follow classical *ApEn* guidelines, although (a) the parameters should be adjusted even for *ApEn* (e.g., [4, 15–20]) and (b) properties of *XApEn* are different and straightforward *ApEn* guidelines might not be proper.

The aim of this paper is to identify the problems that preclude wider *XApEn* implementation and to propose methods for their adjustment. It is proved that the fundamental source of *XApEn* inconsistency is unreliable estimation of

conditional probabilities that are its constitutive elements. It is shown that the parameters required for $XApEn$ are mutually dependent, so the traditional guidelines for their choice need an update. In order to locate the unstable parameter regions, the $XApEn$ values estimated from the artificially simulated data are compared to the true $XApEn$ values calculated by a formula. Following these observations, we derived equations that jointly select stable parameters for reliable $XApEn$ estimation in real signals, for which exact mathematical descriptions are unavailable. The derived equations are tested using a wide variety of cardiovascular signals: using typical short records of ambulatory patients and using long signals recorded from small laboratory animals.

2. Materials and Methods

2.1. Experimental Data. The study used three types of data: artificial, animal, and human. The purpose of artificial data was to test the stability of estimated values, comparing them to the exact values calculated by the derived formulae, and also to develop a new set of threshold equations. The data recorded from animals were also used for method development and for its repeated testing, enabled by the record length. The data recorded from human volunteers were used to validate the derived methods in typical real applications where the records are short.

The animal data include systolic blood pressure (SBP) signals and pulse interval (PI) signals recorded from two groups of rats: the first group comprised outbred male Wistar normotensive rats, while the second group comprised outbred male borderline hypertensive rats. Ten days before the recordings, radiotelemetric probes (TA11PA-C40, DSI, Transoma Medical) were implanted in the abdominal aorta. Blood pressure (BP) waveforms were digitally recorded and transmitted using Dataquest A.R.T. 4.0 equipment (DSI, St. Paul, MN, USA) with sampling frequency $f_s = 1000$ Hz. The systolic blood pressure (SBP) was determined at the receiver as waveform local maxima. The pulse interval was determined as an interval between the successive moments of maximal gradients $(\Delta(BP)/\Delta T)_{\max}$, that is, the local maximal change (Δ) of blood pressure (BP) waveform per sampling interval ΔT . Pulse interval is an acceptable alternative to R-R interval of an electrocardiogram. During the experiments, the animals were exposed to two types of stress: shaker stress, with rats positioned on a platform shaking at 200 cycles/min, and restraint stress, with rats placed in a Plexiglas restrainer tube (ID 5.5 cm with pores) in the supine position, with a detailed description given in [24]. These data were already used for other research purposes [20, 24, 25]. The duration of records in baseline, restraint stress, and shaker stress was 20 minutes, 60 minutes, and 20 minutes, respectively. Such a wide variety of data (different types of rats and different stressing surroundings) ensured the correct validation of the derived methods. Very slow signal component (trend), an outcome of free rat movements, was eliminated using a high-pass filter designed specifically for biomedical signals [26]. The cut-off frequency f_H is tuned to the mean pulse interval of a subject: f_H [Hz] = 0.011/(mean(PI) [s]). Its value ranges from 0.055 Hz to 0.085 Hz in this experiment. The time series

passed the stationarity, so the remaining signals were taken from 12 Wistar normotensive rats (total of 24 signals: 12 baseline, 6 restraint, and 6 shaker) and from 13 borderline hypertensive rats (total of 26 signals: 13 baseline, 7 restraint, and 6 shaker signal sets). All experimental procedures in this study conformed to the European Communities Council Directive of 24 November 1986 (86/609/ECC) and the Guidelines on Animal Experimentation of the School of Medicine, University of Belgrade.

Short data series (number of samples $N \leq 750$) were recorded from 41 medically examined male healthy volunteers positioned in a hospital bed, using Task Force® Monitor with sampling frequency equal to 1000 Hz (CNSystems Medizintechnik AG, Graz, <http://www.cnsystems.com/products/task-force-monitor>). Systolic blood pressure was obtained as local maxima of blood pressure waveforms, while R peaks were found as local maxima of electrocardiogram waveforms. The intervals between the successive R peaks form the R-R interval time series. Sensors were attached by medical professionals and patients were motionless, ensuring the reduced number of artifacts and stationarity of the signals. Short data lengths are typical for ambulatory monitoring, up to 10 minutes, and also ensure tranquility of the patient and stationarity of the recorded data. The experiment followed the ethical standards stated by the School of Medicine, University of Belgrade, and a signed permission was collected from each volunteer.

Artificial data included samples with uniform distributions, generated using a reliable proprietary machine independent random number generator, and the samples with normal distribution generated using the polar method. For each one of the parameter sets, a hundred of the artificial data series were generated, each one comprising up to $N = 5000$ data samples. Program code was written in Matlab 2014b.

An illustration of the time series referenced within this study is shown in Figure 1.

2.2. XApEn in Brief. The procedure for $XApEn$ estimation [22] implies $x_i \in \mathbf{X}$, $i = 1, \dots, N$, as a “master” time series and $y_j \in \mathbf{Y}$, $j = 1, \dots, N$, as its “follower,” where N is the number of signal samples. In biomedical data, SBP is considered as a master signal \mathbf{X} , and PI is considered as its follower \mathbf{Y} . The series are divided into the vectors of length m :

$$\begin{aligned} \mathbf{X}_m^{(i)} &= [x_i, x_{i+\tau}, \dots, x_{i+(m-1)\cdot\tau}], \\ & i = 1, \dots, N - (m-1) \cdot \tau \\ \mathbf{Y}_m^{(j)} &= [y_j, y_{j+\tau}, \dots, y_{j+(m-1)\cdot\tau}], \\ & j = 1, \dots, N - (m-1) \cdot \tau. \end{aligned} \quad (1)$$

Parameter τ can be introduced for vector decorrelation [21] but in most studies it is set to 1 [beat]. A vector from the master series is usually regarded as a “template.”

Distance between two vectors is defined as a maximal absolute sample difference:

$$\begin{aligned} d(\mathbf{X}_m^{(i)}, \mathbf{Y}_m^{(j)}) &= \max_{k=0, \dots, m-1} |x_{i+k\tau} - y_{j+k\tau}|, \\ & i, j = 1, \dots, N - (m-1) \cdot \tau. \end{aligned} \quad (2)$$

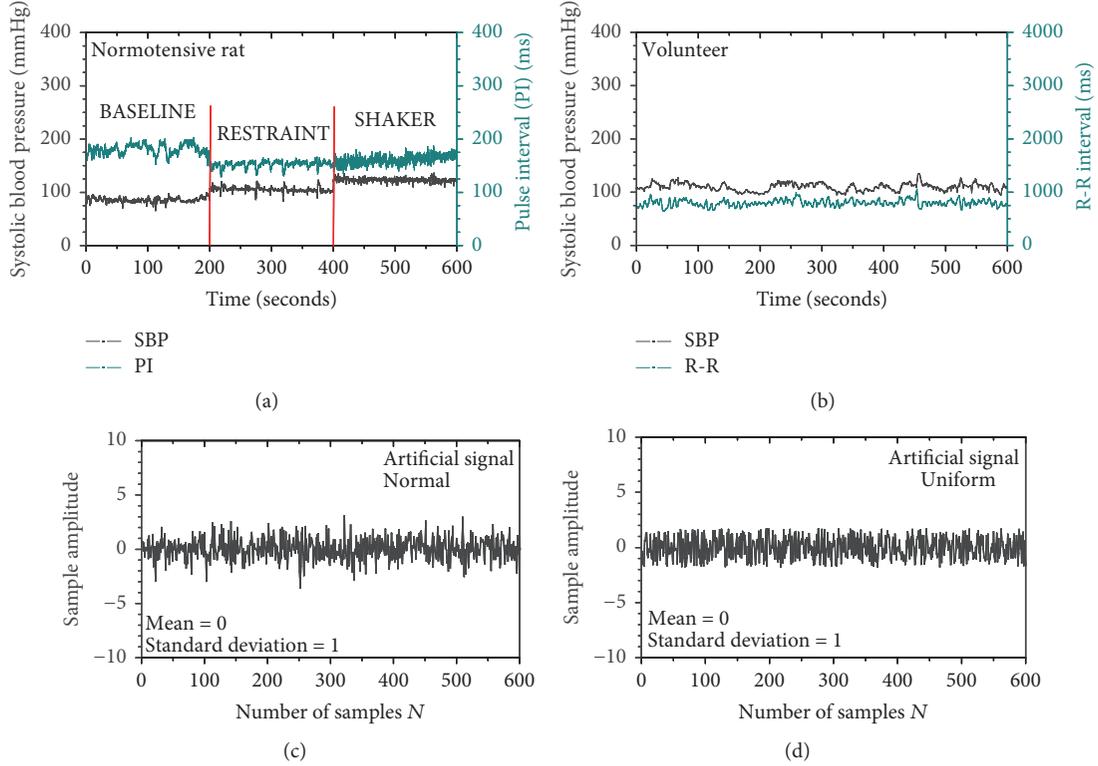


FIGURE 1: Examples of real cardiovascular signals and of artificially generated signals. (a) SBP signal (gray) and PI signal (green) recorded from a rat at baseline condition, exposed to restraint stress and exposed to shaker stress. (b) SBP (gray) and R-R (green) signals recorded from a healthy volunteer. (c, d) Artificially generated signals with zero mean and standard deviation equal to 1; normal distribution (c) and uniform distribution (d).

The vectors are considered as “similar” if their distance is less than or equal to the predefined threshold (tolerance) r . The procedure of finding the similarities is called “template matching.” Standard scoring of time series must be performed; otherwise, vectors from different sources would be incomparable [22]. Standard scoring implies centralization and normalization of series x , $(x - \mu)/\sigma$, where μ denotes mean and σ denotes standard deviation. For a correct estimation of μ and σ , the observed time series must be stationary at least in a wide sense.

A conditional probability that a vector \mathbf{Y} is within the distance r from a particular template $\mathbf{X}_m^{(i)}$, $i = 1, \dots, N - (m - 1) \cdot \tau$, is estimated as follows:

$$\begin{aligned} \hat{p}_i^m(r) &= \Pr \{d(\mathbf{X}_m, \mathbf{Y}_m) \leq r \mid \mathbf{X}_m = \mathbf{X}_m^{(i)}, \mathbf{Y}_m \in \mathbf{Y}, r > 0\} \\ &= \frac{1}{N - (m - 1) \cdot \tau} \sum_{j=0}^{N - (m - 1) \cdot \tau} I \{d(\mathbf{X}_m^{(i)}, \mathbf{Y}_m^{(j)}) \leq r\}, \end{aligned} \quad (3)$$

where “ $\hat{\cdot}$ ” denotes an estimate and $I\{\cdot\}$ denotes an indicator function that is equal to 1 if the condition it indicates is fulfilled (distance less than or equal to r). Denoting the

number of “ones” with $N_1(i)$, the conditional probability (3) can be also expressed as

$$\begin{aligned} \hat{p}_i^m(r) &= \frac{N_1(i)}{N - (m - 1) \cdot \tau}, \\ N_1(i) &= \sum_{j=0}^{N - (m - 1) \cdot \tau} I \{d(\mathbf{X}_m^{(i)}, \mathbf{Y}_m^{(j)}) \leq r\}, \quad i = 1, \dots, N - (m - 1) \cdot \tau. \end{aligned} \quad (4)$$

Probabilities (3) and (4) describe a binary event, as the vector $\mathbf{Y}_m^{(j)}$ can either match or mismatch its template $\mathbf{X}_m^{(i)}$. A precise mathematical derivation of the number of binary events necessary for reliable probability estimation is given in [27]. The derivation in [27] is based assuming a normal approximation of binomial distribution, in our case observing the number of matches found during $N - m + 1$ template matching trials. The resulting tables and graphs point out the probability that the estimated value $\hat{p}_i^m(r)$, for a given number of matching trials, lies within specific boundaries known as “confidence interval.” From the graphically presented results in [27], it can be seen that a 95% confidence interval of about $(0.55 \cdot p_i^m(r), 1.8 \cdot p_i^m(r))$, which is generally considered to be a reasonable uncertainty, is assessed if the number of trials is $10/p_i^m(r)$. So, if the number of tests exceeds $10/p_i^m(r)$, it is considered as a “weak criterion” that yields satisfactory good estimates of binary probabilities [27].

The number of matches that is equal to $100/p_i^m(r)$ is considered as a “strong criterion” for which the boundaries are tighter, with probability estimation less uncertain, but usually it is more difficult to be achieved as it requires long time series. The corresponding confidence intervals, both for weak and for strong criteria, are derived in [27] as follows:

$$\begin{aligned} \Pr \{0.55 \cdot p_i^m(r) \leq \widehat{p}_i^m(r) \leq 1.8 \cdot p_i^m(r)\} &= 0.95 \\ &\text{weak criterion;} \\ \Pr \{0.8 \cdot p_i^m(r) \leq \widehat{p}_i^m(r) \leq 1.25 \cdot p_i^m(r)\} &= 0.95 \\ &\text{strong criterion.} \end{aligned} \quad (5)$$

The next step in *XApEn* estimation is to average logarithms of the probability estimates (4) over all template values and to form a summand $\widehat{\Phi}^{(m)}$:

$$\widehat{\Phi}^{(m)}(r, N, \tau) = \frac{1}{N - (m - 1) \cdot \tau} \sum_{i=0}^{N-(m-1)\tau} \ln(\widehat{p}_i^m(r)). \quad (6)$$

The procedure is repeated for the vectors of length $m + 1$, yielding the estimate:

$$\begin{aligned} \widehat{XApEn}(m, r, N, \tau) &= \widehat{\Phi}^{(m)}(r, N, \tau) \\ &\quad - \widehat{\Phi}^{(m+1)}(r, N, \tau). \end{aligned} \quad (7)$$

Due to the fact that in *XApEn* two different time series are compared, a self-match (matching a template to itself), alleged source of bias in *ApEn*, cannot occur. Unfortunately,

$$\widehat{\Phi}^{(m)}(r, N, \tau) = \begin{cases} -k \cdot \left(1 - \frac{N_0}{(N - (m - 1) \cdot \tau)}\right) & \text{(6) with } \widehat{p}_i^m(r) = 0 \text{ skipped;} \\ -k \cdot \left(1 - \frac{N_0}{(N - (m - 1) \cdot \tau)}\right) - N_0 \cdot \frac{\ln(N - (m - 1) \cdot \tau)}{(N - (m - 1) \cdot \tau)}, & \text{correction (8);} \\ -k & \text{correction (9).} \end{cases} \quad (10)$$

If N tends to infinity, N_0 converges to zero and all $\widehat{\Phi}$ estimates converge to the same true value, that is, to $-k$. But if N is limited, the underestimation of $\widehat{\Phi}$ in (6) (with $\widehat{p}_i^m(r) = 0$ skipped) and (8) is proportional to N_0 . Equation (8) additionally suffers from a nonlinear bias due to the false self-matching, thus repeating the problem of genuine self-matching in *ApEn*. The preferred correction is given by (9) and it was used in further evaluation. However, as it would be clarified later on, if the *XApEn* parameters are properly chosen, the need for this adjustment is reduced.

2.3. Parameter-Induced *XApEn* Inconsistency. The parameters necessary for *XApEn* estimation—time series length N , threshold level r , vector size m , and time lag τ —create a huge working space, but with unstable regions. Recent *ApEn* contributions have reviewed this problem [4, 19], earlier also noticed in [6, 8, 15–18, 20]. But parameter study for *XApEn*

comparison of vectors from different series also means that there is no guarantee that a template of length m in the master time series has at least one match in the follower time series; hence, a lack of matches ($\widehat{p}_i^m(r) = 0$) can occur. The consequence is a logarithm of zero in (6) that makes *XApEn* estimate impossible, pointed out in [5] as well. A usual correction ignores and skips $\widehat{p}_i^m(r) = 0$ cases. Such a correction decreases the numerator in (6) with respect to its denominator and considerably underestimates the summand $\widehat{\Phi}$. We propose two summand $\widehat{\Phi}^{(m)}$ corrections, both different from the corrections in [5]. The first one is analogous to *ApEn* and implements artificial self-matches $\widehat{p}_i^m(r) = 1/(N - (m - 1) \cdot \tau)$ (see (8)). The second correction excludes $\widehat{p}_i^m(r) = 0$ cases from the numerator of (6) (see (9)). If $N_0 = \sum_{i=0}^{N-(m-1)\tau} I\{\widehat{p}_i^m(r) = 0\}$ denotes a total number of zero matches, then

$$\begin{aligned} \widehat{\Phi}_{\text{COR1}}^{(m)}(r, N, \tau) &= \\ &= \frac{\left(\sum_{i=0, \widehat{p}_i^m(r) \neq 0}^{N-(m-1)\tau} \ln(\widehat{p}_i^m(r))\right) - N_0 \cdot \ln(N - (m - 1) \cdot \tau)}{(N - (m - 1) \cdot \tau)} \end{aligned} \quad (8)$$

$$\widehat{\Phi}_{\text{COR2}}^{(m)}(r, N, \tau) = \frac{\sum_{i=0, \widehat{p}_i^m(r) \neq 0}^{N-(m-1)\tau} \ln(\widehat{p}_i^m(r))}{(N - (m - 1) \cdot \tau - N_0)}. \quad (9)$$

A simple abstract experiment analyzes the proposed corrections. Suppose that all nonzero probabilities are equal to an arbitrary constant, $\widehat{p}_i^m(r) = e^{-k}$, $k > 0$. Then, summands $\widehat{\Phi}$ given by (6) with $\widehat{p}_i^m(r) = 0$ skipped and given by (8) and by (9) are equal to

does not exist, except for a brief comment in [5] (to the best of the authors’ knowledge). A few contributions that implement *XApEn* follow the standard guidelines from the original *ApEn/XApEn* papers [1–3, 22]: parameters $m = 1$ and $r = 0.2$ are implemented in [27]; parameters $m = 2$ or 3, $r = 0.15$, and $N = 1000$ are implemented in [28, 29], while parameters $m = 1$ and $r = 0.35$ are used in [30]. It should be noted that the threshold value r in *ApEn* studies is calculated with respect to the standard deviation σ , for example, $r = 0.2 \cdot \sigma$. In *XApEn* studies, standard deviation is by definition set to $\sigma = 1$, due to the obligatory standard scoring [22].

An illustrative example of contradictory results induced by different working points is presented in Figure 2. The contradictory results are noticed in borderline hypertensive rats exposed to shaker stress. Embedded plots show a “*XApEn* threshold profile” [17]: *XApEn* values estimated for a range of threshold r values, keeping all the other parameters fixed.

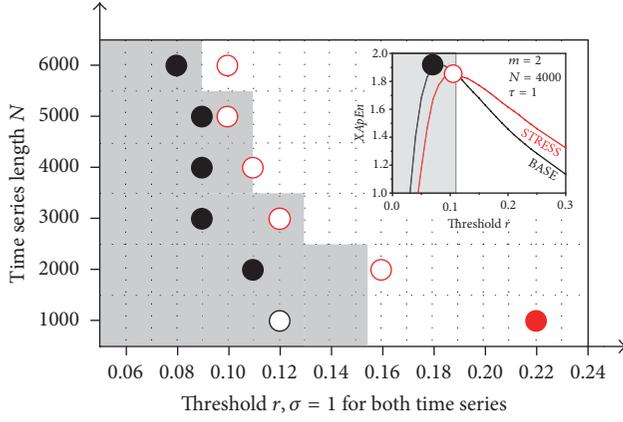


FIGURE 2: Illustrative example of inconsistent results: a threshold-length (r - N) plane for $XApEn$ estimated from SBP and PI signals of rats at baseline (black lines and circles) and stress (red lines and circles). Gray regions: estimated $XApEn$ is higher at baseline than at stress. White regions: $XApEn$ is higher at stress than at baseline. Circles: threshold r_{\max} for which $XApEn$ reaches its maximal value; for a particular N , the empty circle denotes the lower of two maximal $XApEn$ values. Embedded plot shows a complete threshold profile for $N = 4000$. Results are obtained from borderline hypertensive rats exposed to shaker stress.

To obtain a profile, threshold level was gradually increased with increments equal to 0.01 for $r < 0.2$ and equal to 0.05 for $r \geq 0.2$. The black line corresponds to $XApEn$ estimated from SBP and PI signals of rats in baseline conditions, while the red line corresponds to the threshold profile estimated from rats at stress. An intersection of these two lines divides the plot area into two regions: in the “gray” region, $XApEn$ values are higher in baseline conditions than in stress, while in the “white” region $XApEn$ contradictorily becomes lower at baseline than at stress. The maximal value of $XApEn$ (the maximum of the threshold profile line) is denoted by a circle. An “empty” circle denotes the lower one of these two maximums. A complete threshold-length (r - N) plane is shown in the remaining part of Figure 2. A threshold profile is estimated for each particular time series length N . The thresholds positions r_{\max} for which $XApEn$ profile reaches its maximum [17] are shown by circles. We have assumed a time lag $\tau = 1$.

The symbolic presentation in Figure 2 distinguishes three different sources of contradictory results (flip-flop effect):

- For a fixed length N , the $XApEn$ “flips” from being lower during stress for lower threshold r to being higher during stress for higher threshold r ; for example, for $N = 1000$, this change occurs for $r \approx 0.15$.
- For a fixed threshold r , an undesirable change of $XApEn$ from being lower at stress to being higher at stress occurs at particular time series lengths; for example, if $r = 0.12$, this change occurs between $N = 3000$ and $N = 4000$.
- For r_{\max} chosen for a threshold (r_{\max} is a threshold for which $XApEn$ reaches its maximal value, [15–17]), $XApEn$ changes from being higher at stress, if

$N = 1000$, shown by a full red circle, to being lower at stress for $N > 2000$, shown by empty red circles.

2.4. Explicit Formulae for Error Estimation. The constitutive $XApEn$ elements are conditional probabilities $\hat{p}_i^m(r)$ ((3) and (4)), which are subject to an estimation error $\varepsilon_i^m(r)$ each:

$$\begin{aligned} \varepsilon_i^m(r) &= p_i^m(r) - \hat{p}_i^m(r) \implies \\ \hat{p}_i^m(r) &= p_i^m(r) \cdot \left(1 - \frac{\varepsilon_i^m(r)}{p_i^m(r)}\right), \end{aligned} \quad (11)$$

$$i = 1, \dots, N - (m - 1) \cdot \tau.$$

These errors are logarithmically accumulating to form a summand $\widehat{\Phi}^{(m)}(r, N, \tau)$ estimation error $\varepsilon_{\Phi}^m(r, N, \tau)$,

$$\begin{aligned} \widehat{\Phi}^{(m)}(r, N, \tau) &= \frac{1}{N - (m - 1) \cdot \tau} \\ &\cdot \sum_{i=0}^{N-(m-1)\cdot\tau} \left(\ln(p_i^m(r)) + \ln\left(1 - \frac{\varepsilon_i^m(r)}{p_i^m(r)}\right) \right) \implies \end{aligned} \quad (12)$$

$$\varepsilon_{\Phi}^m(r, N, \tau) = \frac{-1}{N - (m - 1) \cdot \tau}$$

$$\cdot \sum_{i=0}^{N-(m-1)\cdot\tau} \ln\left(1 - \frac{\varepsilon_i^m(r)}{p_i^m(r)}\right),$$

and $XApEn$ error $\varepsilon_X^m(r, N, \tau)$,

$$\begin{aligned} \varepsilon_X^m(r, N, \tau) &= \frac{1}{N - m \cdot \tau} \sum_{i=0}^{N-m\cdot\tau} \ln\left(1 - \frac{\varepsilon_i^{m+1}(r)}{p_i^{m+1}(r)}\right) \\ &\quad - \frac{1}{N - (m - 1) \cdot \tau} \sum_{i=0}^{N-(m-1)\cdot\tau} \ln\left(1 - \frac{\varepsilon_i^m(r)}{p_i^m(r)}\right). \end{aligned} \quad (13)$$

To find a stable parameter region where estimation errors would be minimized, the true conditional probabilities $p_i^m(r)$ must be evaluated for each template $\mathbf{X}_{im}^{(i)}$:

$$\begin{aligned} p_i^m(r) &= \prod_{k=0}^{m-1} \int_{x_{i+k\tau-r}}^{x_{i+k\tau+r}} f_y(y) \cdot dy, \\ i &= 1, \dots, N - (m - 1) \cdot \tau, \end{aligned} \quad (14)$$

where $f_y(y)$ denotes the probability density function (p.d.f.) of the signal \mathbf{Y} (follower). A probability $\int_{x_{i+k\tau-r}}^{x_{i+k\tau+r}} f_y(y) \cdot dy$ corresponds to an event where a sample $y_{j+k\tau}$ is within the distance $\pm r$ from a template sample $x_{i+k\tau}$.

Unfortunately, biomedical signals are an outcome of numerous complex physiological inputs and feedback, so exact formulae for their p.d.f. have not been determined yet. Instead, we used artificially generated data with normal and uniform distribution:

$$f_y^{(U)}(y) = \begin{cases} \frac{1}{2 \cdot \sqrt{3}}, & -\sqrt{3} \leq y \leq \sqrt{3} \\ 0, & \text{elsewhere,} \end{cases}$$

uniform p.d.f.

$$f_y^{(N)}(y) = \begin{cases} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-y^2/2}, & \text{normal p.d.f.} \end{cases} \quad (15)$$

The artificial data were used both for the master and for the follower series, generating four experimental data sets: a master and its follower both normal, both uniform, normal master with uniform follower, and uniform master with normal follower.

The corresponding true values of conditional probabilities, assuming $\tau = 1$, are

$$p_i^m(r) = \begin{cases} \frac{\prod_{k=0}^{m-1} (\min(\sqrt{3}, x_{i+k} + r) - \max(-\sqrt{3}, x_{i+k} - r))}{(2 \cdot \sqrt{3})^m}, & |x_{i+k}| - r < \sqrt{3}, \text{ uniform p.d.f.} \\ 0, & |x_{i+k}| - r \geq \sqrt{3} \text{ uniform p.d.f.} \\ \prod_{k=0}^{m-1} \frac{1}{\sqrt{2 \cdot \pi}} \int_{x_{i+k}-r}^{x_{i+k}+r} e^{-y^2/2} \cdot dy = \prod_{k=0}^{m-1} (\varphi(x_{i+k} + r) - \varphi(x_{i+k} - r)) & \text{normal p.d.f.} \end{cases} \quad (16)$$

Normal cumulative distribution function $(1/\sqrt{2 \cdot \pi}) \int_{-\infty}^z e^{-y^2/2} \cdot dy$ in the lower part of (16) is denoted by $\varphi(z)$ instead of $\Phi(z)$, to avoid confusion with the summand Φ . With (16) known, all the relevant estimation errors (11)–(13) can be found.

2.5. Reference Thresholds for Cardiovascular Signals. The error analysis in artificial data has a privilege to evaluate a reference point (true value) using an exact formula. The problem with natural signals is that their data samples are statistically dependent. Besides, the signal nature is not completely revealed and there are no mathematical expressions to describe it. So, there exists no calculated reference value to which the estimated results could be compared. The first auxiliary point of reference considering *ApEn* was empirically obtained in [17], where the authors pointed out the existence of *ApEn* maximum and the corresponding threshold r_{\max} for which this maximum is reached. r_{\max} is found by performing a tedious task of estimating *ApEn* for each threshold value separately, thus obtaining a threshold profile. In [15], tedious estimation is substituted by a set of expressions that automatically evaluate a theoretical threshold $r_{\text{TH-A}}$ that is (almost) equal to r_{\max} . In [16], the same authors proposed an alternative set of formulae for $r_{\text{TH-A}}$. It was shown [15, 16] that the empirical threshold level r_{\max} for which maximal value of *ApEn* is obtained when plotting the threshold profile depends on the time series length N , on standard deviations σ of time series $x(i)$, and on standard deviations σ_D of lagged differential signal $(x(i) - x(i - \tau))$, where τ is a time lag (delay) between the samples, $\tau = 1$. The observed dependency was shown to be linear with respect to the square root of σ_D and σ ratio, so the authors performed a computer search for the coefficients that yield the best fit in the least-squares sense to obtain (18). We have adopted the same method to find the $m = 1$ case of a theoretical threshold $r_{\text{TH-A}}$ in (18), necessary for our evaluations, but not covered in [15, 16].

A method for automatic evaluation of auxiliary reference point does not exist for *XApEn*. We observed [25] that the threshold $r_{\text{TH-A}}$ of master time series influences the threshold r_{\max} of *XApEn* and that their difference is linearly related to the square root of $\sigma_{DX} + \sigma_{DY}$, where σ_{DX} and σ_{DY} are standard deviations of the lagged differential series $(x(i) - x(i - \tau))$ and $(y(i) - y(i - \tau))$ [25]. Then, for each artificially created master **X**-follower **Y** combination (normal-normal, normal-uniform, uniform-normal, and uniform-uniform), we generated 100 normalized signal pairs, in total 400 signal pairs per parameter set. For each signal pair, we found the maximum of the threshold profile and the corresponding r_{\max} , and we also estimated σ_{DX} and σ_{DY} . Then, the least-squares fitting is performed with respect to r_{\max} , thus obtaining particular coefficients $a(m)$, $b(m)$, and $c(m)$. The procedure is repeated for different N values, from 300 to 5000 with an increment equal to 100. The results were averaged for each particular $m \in \{1, 2, 3, 4\}$ giving the coefficient sets $a(m)$, $b(m)$, and $c(m)$. Since several sets had almost the same averaged least-squares error, ten out of 25 baseline signals from rats were randomly selected for justification, and the coefficient set with minimal absolute error (averaged over all values of N and m) was selected for a threshold $r_{\text{TH-X}}$, for which *XApEn* approaches its maximal value:

$$\begin{aligned} r_{\text{TH-X}}(m) &= r_{\text{TH-A}}(m) \\ &+ \left| a(m) + \frac{b(m) + c(m) \cdot \sqrt{(\sigma_{DX} + \sigma_{DY})/2}}{\sqrt[4]{N/1000}} \right|, \end{aligned} \quad (17)$$

$$(a, b, c) = \begin{cases} (0, -0.015, 0.03), & m = 1 \\ (-0.02, 0, 0.023), & m = 2 \\ (0, -0.006, 0.043), & m = 3 \\ (0, -0.11, 0.13), & m = 4, \end{cases}$$

where $r_{\text{TH-A}}$ is evaluated from the master time series according to

$$r_{\text{TH-A}}(m) = \frac{e(m) + f(m) \cdot \sqrt{\sigma_{DX}/\sigma_X}}{\sqrt[4]{N/1000}},$$

$$(e, f) = \begin{cases} (-0.01, 0.05), & m = 1 \\ (-0.02, 0.23), & m = 2 \\ (-0.06, 0.43), & m = 3 \\ (-0.11, 0.65), & m = 4. \end{cases} \quad (18)$$

Here, σ_X is standard deviation of a time series \mathbf{X} (in *XApEn* it is always equal to one) while σ_{DX} and σ_{DY} are standard deviations of the lagged differential series ($x(i) - x(i-\tau)$) and ($y(i) - y(i-\tau)$). The expression for $m = 1$ in (18) is presented for the first time in this study, while the expressions for $m = 2, 3, 4$ are taken from [15]. As it is shown in Table 3, the expressions from [15] yield a lower amount of relative error for cardiovascular data used in this study, so they are more appropriate than the similar expressions derived in [16]. The subset of thresholds (17) for $m = 2$ and $m = 3$ was presented in [25].

The threshold $r_{\text{TH-X}}$ is only halfway towards the reliable working point. As it would be shown, if threshold $r_{\text{TH-A}}$ is applied, the reliability criteria stated by $N > 10/p_i^m(r)$ (or $100/\hat{p}_i^m(r)$) are not satisfied. A different interpretation of these criteria states that at least 10 (or 100) events should be found, in order to have the estimated event probability fitted within the confidence bounds (5) with the certainty of 95%.

Obviously, increased m reduces the matching probability of the template and its follower, and thus a threshold increase had to be proportional to it. The coefficients in (19) and (20) are obtained using the same data sets as for (17). In this case, the conditional probabilities were estimated, and the fitting was performed among the coefficients that yield less than 5% of unreliable probabilities, over all values of m and over all data lengths:

$$r_{\text{XW}}(m) = 0.2 \cdot (m + 1) + r_{\text{TH-X}}(m),$$

$$m \leq 4, N \leq 500,$$

$$r_{\text{XW}}(m) = 0.17 \cdot (m + 1) + r_{\text{TH-X}}(m),$$

$$m \leq 4, 500 < N \leq 2000, \quad (19)$$

$$r_{\text{XW}}(m) = 0.15 \cdot (m + 1) + r_{\text{TH-X}}(m),$$

$$m \leq 4, N > 2000.$$

In the case of strong reliability criteria, a relationship of threshold (20) with respect to vector length m is nonlinear:

$$r_{\text{XS}}(m) = \left(m + 1 + \frac{10^5}{N^2} \right) \cdot \left(r_{\text{TH-X}}(m) + \frac{(5-m)}{10} \right), \quad (20)$$

$$m \leq 4.$$

For the sake of clarity, the different threshold values are summarized in Table 1.

TABLE 1: Threshold abbreviations and definitions.

Abbreviation	Definition
$r \cdot \sigma$	<i>ApEn</i> threshold, expressed in units of time series
r	<i>XApEn</i> threshold, dimensionless
r_{max}	Threshold obtained from threshold profile, for which (<i>X</i>) <i>ApEn</i> reaches its maximal value [17]
$r_{\text{TH-A}}$	Automatically calculated <i>ApEn</i> threshold using (18)
$r_{\text{TH-X}}$	Automatically calculated <i>XApEn</i> threshold using (17)
r_{XW}	Automatically calculated <i>XApEn</i> threshold for which at least 95% of the estimated probabilities satisfy weak criterion.
r_{XS}	Automatically calculated <i>XApEn</i> threshold for which at least 95% of the estimated probabilities satisfy strong criterion.

2.6. Reliability Criteria in Statistically Dependent Data. The confidence bounds (5) hold for statistically independent binary events [27]. Template matching in artificial data satisfies this constraint: observing an $m = 1$ template $[x_k]$, its successive matching tests $|x_k - y_i| \leq r$ and $|x_k - y_{i+1}| \leq r$ at the respective positions i and $i + 1$ are independent, as an absolute difference of a constant x_k and independent samples y_i and y_{i+1} does not destroy independency. A similar line of reasoning holds if $m > 1$. It follows that the estimated probabilities $\hat{p}_i^m(r)$ have a scaled binomial distribution [27], so the reliability criteria and confidence intervals (5) can be applied. If the data were statistically dependent, the only change is an increase of confidence bounds, without affecting the estimated values or the reliability criteria [27]. However, true confidence bounds in real biomedical data cannot be defined as true $p_i^m(r)$ is unknown, so statistical dependency is of no importance for real data. The reliability still can be tested in a form of $N - m + 1 > (10 \text{ or } 100)/\hat{p}_i^m(r)$, showing whether the obtained value $\hat{p}_i^m(r)$ is estimated from a sufficient amount of data.

3. Results and Discussion

3.1. Estimation Errors in Artificial Data. For *XApEn* error estimation, master series \mathbf{X} and a hundred of its followers \mathbf{Y} were artificially generated for a range of $(r, N, m, (\tau = 1))$ parameters, and the same values exactly were calculated using the derived formulae. The results for all four data sets, due to the size, are presented within the Supplementary Material (available online at <https://doi.org/10.1155/2017/8365685>).

Relative conditional probability errors $|\varepsilon_i^m(r)|/p_i^m(r)$ [%] are plotted in Figure 3, for series \mathbf{X} and \mathbf{Y} both normally distributed and for $\tau = 1$. Satisfactory results are obtained for an extremely atypical threshold level $r = 1.50$. For a typical threshold $r = 0.15$, error is below 10% only for large N and $m = 1$. The results are in accordance with reliability criteria $N - m + 1 > (10 \text{ or } 100)/p_i^m(r)$ stating that N or $p_i^m(r)$ should be large. Indeed, $p_i^m(r)$ is defined as a product of m probabilities (14), so it is at its largest if $m = 1$, while increased threshold r augments the number of template matches $N_1(i)$ (4), again enlarging $p_i^m(r)$.

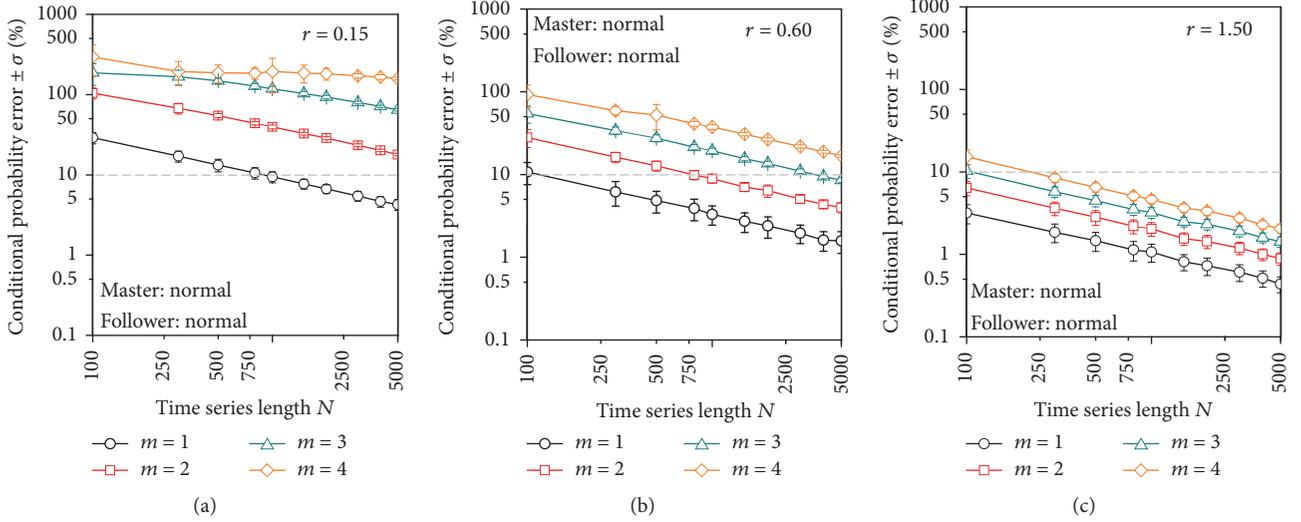


FIGURE 3: The relative conditional probability error [%] in artificial data for a typical threshold value $r = 0.15$ (a), for a higher threshold $r = 0.60$ (b), and for an unusually high threshold $r = 1.50$ (c); the gray dashed horizontal line marks 10% error; results are presented in a log-log plane as mean $\pm \sigma$ (standard deviation). Artificial master and follower series are both normally distributed.

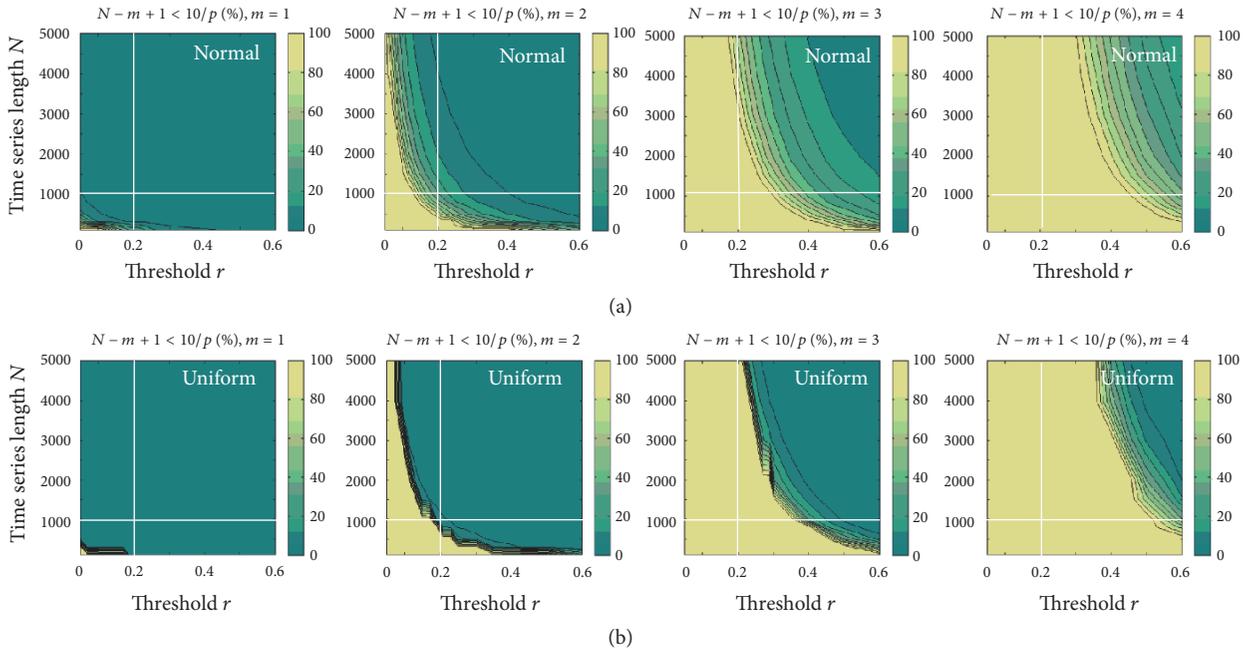


FIGURE 4: Percentage of estimates $\hat{p}_i^m(r)$ in artificial data that fail to satisfy the criterion $N - m + 1 \geq 10/\hat{p}_i^m(r)$ [%], presented in threshold-length (r - N) plane, for $m = 1, 2, 3, 4$; failures are presented by color changes from green (0% failures) to yellow (100% failures); (a) X and Y normally distributed; (b) X and Y uniformly distributed; white horizontal and vertical lines cut each plane at typical parameter values $r = 0.2$ and $N = 1000$.

The reliability of $\hat{p}_i^m(r)$ estimate is checked comparing the data length $N - m + 1$ with the required minimum $10/\hat{p}_i^m(r)$. The percentage of test failures is presented in Figure 4, in an r - N plane, where 0% failures are presented by green color that gradually changes to yellow (100% failures). The white line marks classical r and N values, revealing that, for $m = 3$ and 4, the estimates that implement classical parameters are unreliable. Nonlinear interrelation of r , N , and m in Figure 4 also points out that the choice of parameters that would yield a stable working point is not a straightforward task.

The estimation error of summand $\hat{\Phi}$ (12) exposes two problems. The first one is a lack of matches ($\hat{p}_i^m(r) = 0$) that induce logarithms of zero in (6) which is, as shown in Figure 5, a subset of all unreliable probabilities from Figure 4. The second problem is an error accumulation (12).

Figure 6 shows a line graph of summand Φ and its changes with respect to series length N if it is evaluated by a formula (exact values), if it is estimated in a traditional way, and if a correction (9) is applied. This correction does attenuate estimation errors, even for high m values, but it

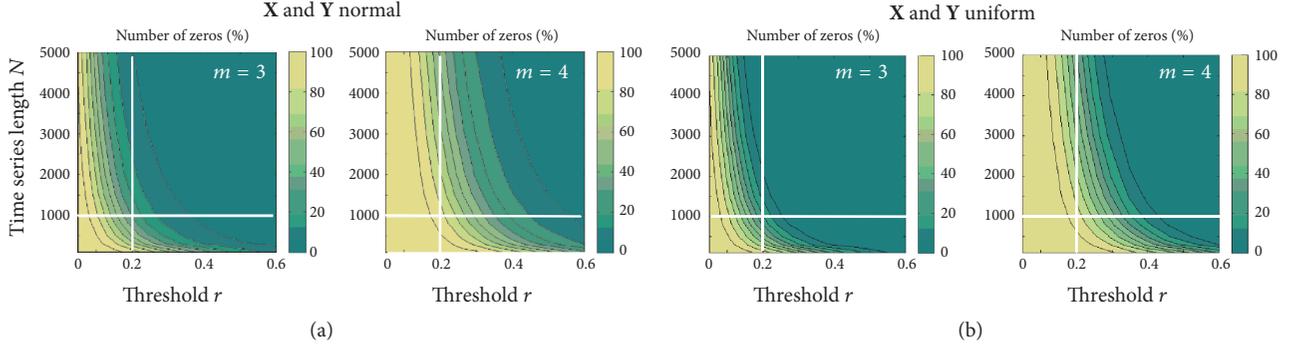


FIGURE 5: Number of $\hat{p}_i^m(r) = 0$ estimates [%] in artificial data presented in threshold-length (r - N) plane for $m = 3$ and 4; color changes from green (0% zero matches) to yellow (100% zero matches); (a) X and Y normally distributed; (b) X and Y uniformly distributed.

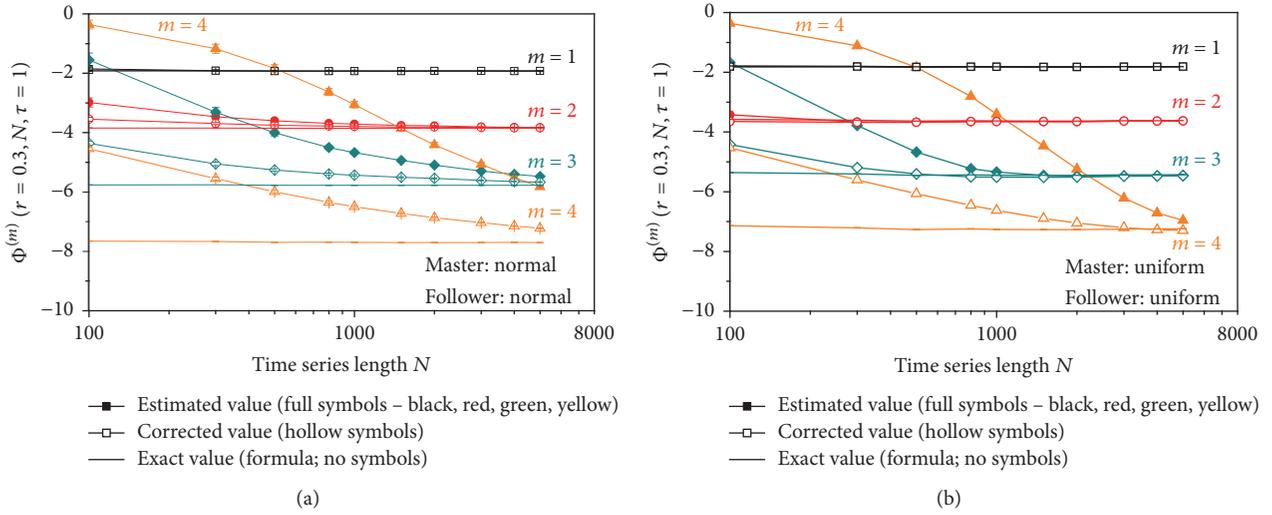


FIGURE 6: Summand Φ for $r = 0.3$: estimated omitting the zero probability, estimated and corrected, and calculated exactly using a formula; (a) X and Y normally distributed; (b) X and Y uniformly distributed.

cannot be used alone, as it removes a subset of the problem only. The convergence of summand $\widehat{\Phi}$ towards its true value depends on data type: uniform data achieve the correct limit for shorter lengths N . The convergence is also m -dependent, requiring different data lengths to achieve the same limit. Since \widehat{XApEn} is defined as a difference of $\widehat{\Phi}$ values (7), the fact that each summand $\widehat{\Phi}$ adds its own and independent estimation error is a further source of \widehat{XApEn} instability.

Figure 7 shows two-dimensional $XApEn$ profiles. Values calculated using the derived formulae ((a) and gray lines in (b) and (c)) verify the known fact that $XApEn$ linearly decreases with the logarithm of threshold value [5]. The major conclusion is that the true $XApEn$ values are influenced neither by time series length N nor by template length m . Frequent notions that $XApEn$ monotonously increases with N until a plateau is reached or that its value decreases with m are a consequence of probability underestimation that decreases with series length and eventually disappears. The panels in the second row of (a) show corrected \widehat{XApEn} estimates that are in good accordance with true values for $m = 1$, but showing increased discrepancies for $m = 2$ and

3. If no correction is applied (the third row of (a)), \widehat{XApEn} substantially diverges from the correct values. (b) and (c) show \widehat{XApEn} as a line graph. The discrepancy between the true and estimated values points out that generally accepted guidelines for parameter choice do not offer a safe \widehat{XApEn} estimation.

3.2. Stable Working Point in Cardiovascular Environment. To verify the accuracy of the derived equation (17), we followed the recommendations from [15, 16], where their equations are validated with respect to $ApEn$ relative error. For each one of SBP-PI (or R-R) and PI (or R-R)-SBP signal pairs, of animals (or volunteers), a threshold r_{TH-X} was evaluated according to (17) and used as a parameter for $XApEn(r_{TH-X})$ estimation. The true maximal $XApEn(r_{max})$ value is provided by a time-consuming procedure of finding a threshold profile. The corresponding relative error $\epsilon_{XA,max}$ is presented in Table 2. The results estimated from signals recorded in human volunteers are presented in the italic part of Table 2. The relative error is below 2%, except for very short data lengths ($N = 300$) and $m = 1$ when it reaches 6%; unfortunately,

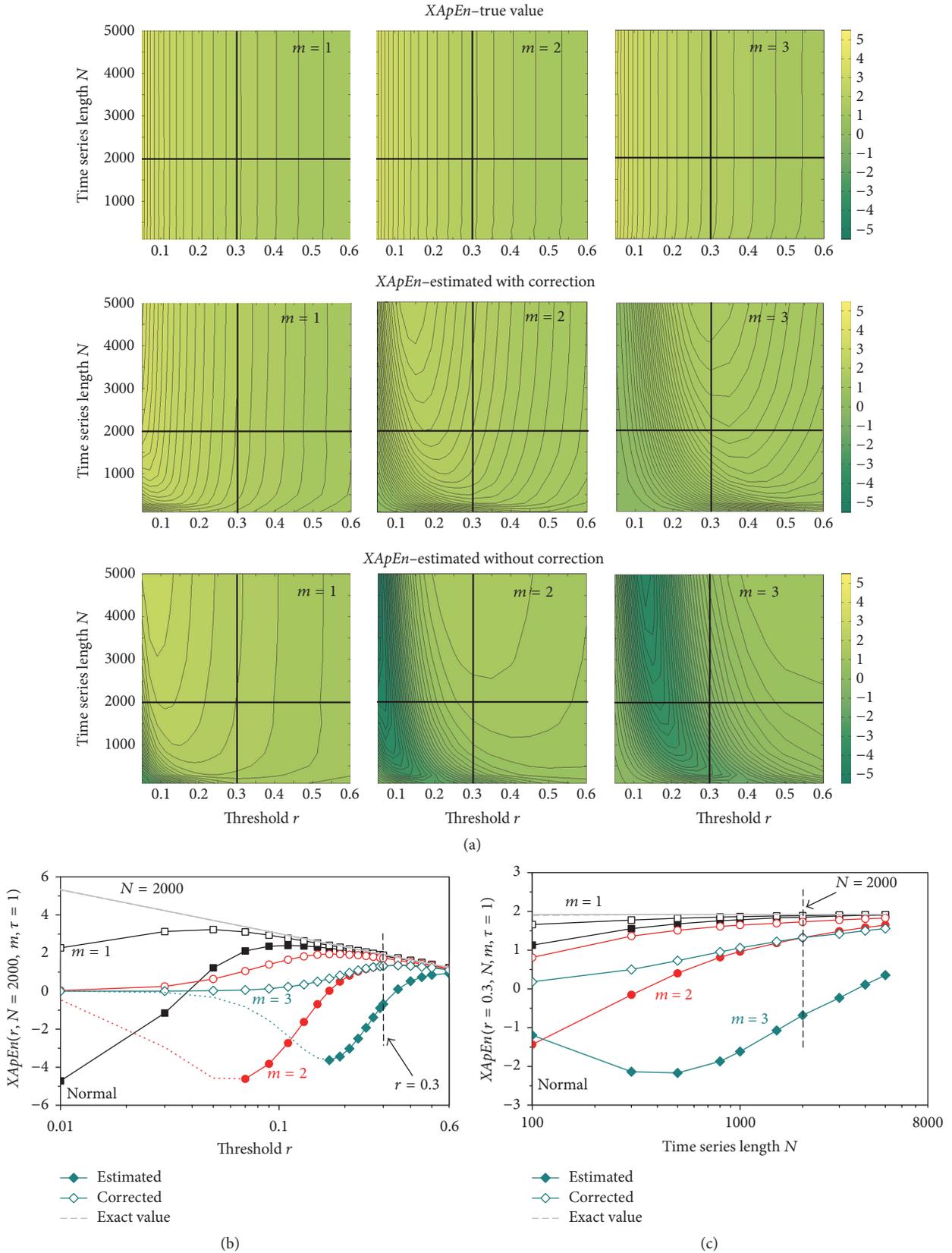


FIGURE 7: $XApEn$ of artificial data with normal distribution in threshold-length (r - N) plane, for $m = 1, 2$, and 3 . The three rows of (a) show, respectively, calculated (true) $XApEn$, $XApEn$ estimated with correction, and $XApEn$ estimated omitting $\log(0)$. Dark green color corresponds to $XApEn$ estimated within the unstable N - r - m parameter region. Light green indicates low $XApEn$, while yellow color indicates high $XApEn$. Horizontal cuts at $N = 2000$ are presented as line graph in (b); vertical cuts at $r = 0.3$ are presented as line graph in (c).

TABLE 2: Relative error of $XApEn(r_{TH-X})$ with respect to $XApEn(r_{max})$ in real signals [%].

N	$\varepsilon_{XA,max} = (XApEn(r_{TH-X}) - XApEn(r_{max}) /XApEn(r_{max})) \cdot 100$							
	X = PI, Y = SBP				X = SBP, Y = PI			
	m = 1	m = 2	m = 3	m = 4	m = 1	m = 2	m = 3	m = 4
300	5.16 ± 0.40	1.25 ± 0.28	1.61 ± 0.21	5.51 ± 0.75	6.40 ± 0.61	2.42 ± 0.49	2.85 ± 0.65	5.60 ± 0.88
500	2.57 ± 0.25	1.13 ± 0.21	2.61 ± 0.51	2.91 ± 0.57	3.61 ± 0.39	1.74 ± 0.29	1.68 ± 0.34	2.83 ± 0.60
750	1.57 ± 0.18	0.90 ± 0.21	1.93 ± 0.49	2.06 ± 0.44	2.09 ± 0.26	1.09 ± 0.25	1.56 ± 0.28	1.62 ± 0.43
2000	0.64 ± 0.17	1.07 ± 0.26	1.50 ± 0.45	1.83 ± 0.54	0.49 ± 0.16	1.43 ± 0.29	1.96 ± 1.40	2.62 ± 0.47
2500	0.46 ± 0.07	0.88 ± 0.14	1.29 ± 0.29	1.61 ± 0.42	0.32 ± 0.05	1.01 ± 0.22	1.45 ± 0.29	2.99 ± 0.47
3000	0.80 ± 0.16	0.91 ± 0.18	0.99 ± 0.25	1.22 ± 0.26	0.33 ± 0.08	0.99 ± 0.18	0.94 ± 0.16	2.03 ± 0.38
3500	0.67 ± 0.12	1.04 ± 0.19	1.89 ± 0.67	2.68 ± 1.28	0.52 ± 0.14	1.41 ± 0.33	1.58 ± 0.47	2.64 ± 0.58
4000	1.03 ± 0.20	1.65 ± 0.28	1.58 ± 0.26	1.15 ± 0.23	1.08 ± 0.25	1.22 ± 0.23	1.09 ± 0.24	1.79 ± 0.45
4500	0.68 ± 0.18	1.41 ± 0.39	1.72 ± 0.59	2.80 ± 0.73	0.94 ± 0.18	1.20 ± 0.30	1.90 ± 0.92	2.28 ± 1.33
5000	1.59 ± 0.36	1.57 ± 0.44	1.73 ± 0.61	2.26 ± 0.75	0.42 ± 0.08	2.20 ± 0.39	1.74 ± 0.56	1.64 ± 0.40

Results are presented as mean ± SEM. Results of 41 human volunteers are shown in the first three italic rows. Results of 25 rats in different conditions (total of 50 signals) are presented in roman rows.

TABLE 3: Relative error of $ApEn(r_{TH-A})$ with respect to $ApEn(r_{max})$ [%] in real signals.

N	$\varepsilon_{A,max} = (ApEn(r_{TH-A}) - ApEn(r_{max}) /ApEn(r_{max})) \cdot 100$			
	X = PI		X = SBP	
	r_{TH-A} [15]	r_{TH-A} [16]	r_{TH-A} [15]	r_{TH-A} [16]
2000	3.46 ± 0.53	4.07 ± 0.54	1.95 ± 0.33	2.33 ± 0.35
3000	3.23 ± 0.54	3.82 ± 0.56	2.30 ± 0.35	2.91 ± 0.48
4000	3.72 ± 0.65	4.40 ± 0.67	2.86 ± 0.48	3.33 ± 0.51
5000	3.75 ± 0.67	6.45 ± 1.26	2.68 ± 0.43	5.63 ± 1.29

Error is evaluated according to the guidelines from [15] and from [16]. Results are presented as mean ± SEM, averaged over 100 time series: 50 data series of Wistar and borderline hypertensive rats in baseline and in stress, repeated for $m = 2$ and for $m = 3$.

this corresponds to the ambulatory records that cannot be too long.

For the sake of comparison, we estimated $ApEn$ error from our data as well. We applied r_{TH-A} calculated according to [15] (see (18)) and according to [16]. The r_{TH-A} relative error is shown in Table 3, providing the following conclusions: (1) error level is lower if the equations from [15] are used and (2) comparison of results in Tables 2 and 3 verifies that our expressions (17) are equal in performance to expressions (18) derived by Lu et al. in [15, 16].

The value r_{TH-X} cannot be applied, since the percentage of the probabilities that fail the weak and strong criteria is almost 100% (Figure 8). But this threshold reflects the nature of data and it can be used as a reference for thresholds r_{XW} and r_{XS} . Their implementation reduces the failures, as shown in Figure 8.

The actual threshold values calculated from real signals are shown in Table 4. For the sake of comparison, threshold values for normally distributed artificial data are also plotted (Figure 9). These thresholds show that the traditional guidelines (0.15, 0.2, and 0.35) considerably underestimate the actual requirements for the reliable $XApEn$ estimation. Even if $m = 1$, the threshold required for strong reliability may exceed $r = 3$. Although it was speculated [1] that detailed system information would be lost if r values were larger than $0.1 \cdot \sigma$ or $0.2 \cdot \sigma$ (for $ApEn$), meaning that if

the threshold exceeds these values a template would always match its follower, in investigated cardiovascular signals, this assumption does not hold as the evaluated high threshold levels are considerably below the amplitude range of the observed data.

The developed sets of threshold expressions are primarily intended for cardiovascular signal analysis. The artificial signals used as well are uniform on segment (with bounded amplitudes) and Gaussian (with increased incidence of small amplitudes, but unbounded). Both artificial distributions are symmetric. The amplitudes of real signals (PI, SBP, and R-R) are bounded by physiological constraints. Signal asymmetry is expressed by skewness factor, shown in Table 5. Cardiovascular data can be skewed to the left and to the right, so Table 5 presents the average of absolute skewness values. SBP and PI signals exhibit a level of asymmetry comparable to Rayleigh distribution (skewness ≈ 0.631), but below the skewness of exponential distribution (skewness ≈ 2). Threshold evaluation in signals with highly skewed distribution would be a subject for follow-up studies. In particular, it would be precious to explore an important class of interspike interval time series, with distributions that are close to exponential. A memoryless property of exponential distribution might affect the values of differential time series standard deviation σ_D , necessary for (17) and (18).

TABLE 4: Threshold values obtained from real signals.

(a) Threshold r_{TH-X} for which $XApEn$ reaches its maximal value (17)

N	$X = PI, Y = SBP$				$X = SBP, Y = PI$			
	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
300	<i>0.05 ± 0.01</i>	<i>0.24 ± 0.01</i>	<i>0.44 ± 0.01</i>	<i>0.62 ± 0.01</i>	<i>0.05 ± 0.00</i>	<i>0.22 ± 0.01</i>	<i>0.40 ± 0.01</i>	<i>0.55 ± 0.02</i>
500	<i>0.05 ± 0.01</i>	<i>0.21 ± 0.00</i>	<i>0.39 ± 0.01</i>	<i>0.53 ± 0.00</i>	<i>0.04 ± 0.00</i>	<i>0.19 ± 0.00</i>	<i>0.35 ± 0.01</i>	<i>0.47 ± 0.01</i>
750	<i>0.04 ± 0.01</i>	<i>0.18 ± 0.04</i>	<i>0.35 ± 0.01</i>	<i>0.47 ± 0.01</i>	<i>0.04 ± 0.01</i>	<i>0.16 ± 0.00</i>	<i>0.31 ± 0.01</i>	<i>0.41 ± 0.01</i>
2000	<i>0.05 ± 0.01</i>	<i>0.19 ± 0.00</i>	<i>0.37 ± 0.01</i>	<i>0.51 ± 0.01</i>	<i>0.04 ± 0.00</i>	<i>0.14 ± 0.00</i>	<i>0.27 ± 0.00</i>	<i>0.60 ± 0.01</i>
2500	<i>0.05 ± 0.01</i>	<i>0.18 ± 0.00</i>	<i>0.35 ± 0.01</i>	<i>0.48 ± 0.01</i>	<i>0.04 ± 0.00</i>	<i>0.13 ± 0.00</i>	<i>0.25 ± 0.01</i>	<i>0.34 ± 0.01</i>
3000	<i>0.04 ± 0.01</i>	<i>0.18 ± 0.00</i>	<i>0.33 ± 0.01</i>	<i>0.46 ± 0.01</i>	<i>0.03 ± 0.00</i>	<i>0.13 ± 0.00</i>	<i>0.24 ± 0.00</i>	<i>0.33 ± 0.05</i>
3500	<i>0.04 ± 0.01</i>	<i>0.17 ± 0.00</i>	<i>0.32 ± 0.01</i>	<i>0.45 ± 0.01</i>	<i>0.03 ± 0.00</i>	<i>0.12 ± 0.00</i>	<i>0.23 ± 0.00</i>	<i>0.31 ± 0.05</i>
4000	<i>0.04 ± 0.01</i>	<i>0.17 ± 0.00</i>	<i>0.31 ± 0.01</i>	<i>0.44 ± 0.01</i>	<i>0.03 ± 0.00</i>	<i>0.17 ± 0.00</i>	<i>0.22 ± 0.00</i>	<i>0.30 ± 0.05</i>
4500	<i>0.04 ± 0.01</i>	<i>0.11 ± 0.00</i>	<i>0.30 ± 0.01</i>	<i>0.43 ± 0.01</i>	<i>0.03 ± 0.00</i>	<i>0.11 ± 0.00</i>	<i>0.21 ± 0.00</i>	<i>0.29 ± 0.01</i>
5000	<i>0.04 ± 0.01</i>	<i>0.16 ± 0.00</i>	<i>0.29 ± 0.01</i>	<i>0.42 ± 0.01</i>	<i>0.03 ± 0.00</i>	<i>0.11 ± 0.00</i>	<i>0.20 ± 0.00</i>	<i>0.29 ± 0.01</i>

(b) Threshold r_{XW} for which at least 95% of the probabilities satisfy $N - m + 1 > 10/\hat{p}_i^m(r_{XW})$ (19)

N	$X = PI, Y = SBP$				$X = SBP, Y = PI$			
	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
300	<i>0.46 ± 0.00</i>	<i>0.84 ± 0.01</i>	<i>1.24 ± 0.01</i>	<i>1.62 ± 0.01</i>	<i>0.45 ± 0.00</i>	<i>0.82 ± 0.01</i>	<i>1.20 ± 0.01</i>	<i>1.55 ± 0.02</i>
500	<i>0.45 ± 0.00</i>	<i>0.81 ± 0.00</i>	<i>1.19 ± 0.01</i>	<i>1.53 ± 0.01</i>	<i>0.44 ± 0.00</i>	<i>0.78 ± 0.00</i>	<i>1.15 ± 0.01</i>	<i>1.47 ± 0.01</i>
750	<i>0.38 ± 0.00</i>	<i>0.69 ± 0.00</i>	<i>1.03 ± 0.01</i>	<i>1.32 ± 0.01</i>	<i>0.38 ± 0.00</i>	<i>0.67 ± 0.00</i>	<i>0.99 ± 0.01</i>	<i>1.26 ± 0.01</i>
2000	<i>0.39 ± 0.00</i>	<i>0.70 ± 0.00</i>	<i>1.05 ± 0.01</i>	<i>1.36 ± 0.01</i>	<i>0.37 ± 0.00</i>	<i>0.65 ± 0.00</i>	<i>0.95 ± 0.01</i>	<i>1.21 ± 0.01</i>
2500	<i>0.35 ± 0.00</i>	<i>0.63 ± 0.00</i>	<i>0.95 ± 0.01</i>	<i>1.23 ± 0.01</i>	<i>0.34 ± 0.00</i>	<i>0.58 ± 0.00</i>	<i>0.85 ± 0.01</i>	<i>1.09 ± 0.01</i>
3000	<i>0.34 ± 0.00</i>	<i>0.63 ± 0.00</i>	<i>0.93 ± 0.01</i>	<i>1.22 ± 0.01</i>	<i>0.33 ± 0.00</i>	<i>0.59 ± 0.00</i>	<i>0.84 ± 0.00</i>	<i>1.08 ± 0.01</i>
3500	<i>0.34 ± 0.00</i>	<i>0.62 ± 0.00</i>	<i>0.92 ± 0.01</i>	<i>1.20 ± 0.01</i>	<i>0.33 ± 0.00</i>	<i>0.57 ± 0.00</i>	<i>0.83 ± 0.00</i>	<i>1.06 ± 0.01</i>
4000	<i>0.34 ± 0.00</i>	<i>0.62 ± 0.00</i>	<i>0.91 ± 0.01</i>	<i>1.19 ± 0.01</i>	<i>0.33 ± 0.00</i>	<i>0.57 ± 0.00</i>	<i>0.82 ± 0.00</i>	<i>1.05 ± 0.01</i>
4500	<i>0.34 ± 0.00</i>	<i>0.61 ± 0.00</i>	<i>0.90 ± 0.01</i>	<i>1.18 ± 0.01</i>	<i>0.33 ± 0.00</i>	<i>0.56 ± 0.00</i>	<i>0.81 ± 0.00</i>	<i>1.04 ± 0.01</i>
5000	<i>0.34 ± 0.00</i>	<i>0.61 ± 0.00</i>	<i>0.89 ± 0.01</i>	<i>1.17 ± 0.01</i>	<i>0.33 ± 0.00</i>	<i>0.56 ± 0.00</i>	<i>0.80 ± 0.00</i>	<i>1.04 ± 0.01</i>

(c) Threshold r_{XS} for which at least 95% of the probabilities satisfy $N - m + 1 > 100/\hat{p}_i^m(r_{XS})$ (20)

N	$X = PI, Y = SBP$				$X = SBP, Y = PI$			
	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
300	<i>1.42 ± 0.00</i>	<i>2.21 ± 0.02</i>	<i>3.28 ± 0.04</i>	<i>4.37 ± 0.09</i>	<i>1.40 ± 0.01</i>	<i>2.12 ± 0.02</i>	<i>3.07 ± 0.05</i>	<i>3.99 ± 0.10</i>
500	<i>1.08 ± 0.00</i>	<i>1.73 ± 0.14</i>	<i>2.60 ± 0.04</i>	<i>3.41 ± 0.01</i>	<i>1.06 ± 0.00</i>	<i>1.65 ± 0.01</i>	<i>2.41 ± 0.00</i>	<i>3.06 ± 0.07</i>
750	<i>0.96 ± 0.00</i>	<i>1.54 ± 0.01</i>	<i>2.28 ± 0.03</i>	<i>2.93 ± 0.05</i>	<i>0.96 ± 0.00</i>	<i>1.47 ± 0.01</i>	<i>2.13 ± 0.03</i>	<i>2.64 ± 0.04</i>
2000	<i>0.91 ± 0.00</i>	<i>1.50 ± 0.01</i>	<i>2.29 ± 0.03</i>	<i>3.06 ± 0.05</i>	<i>0.88 ± 0.00</i>	<i>1.33 ± 0.01</i>	<i>1.89 ± 0.02</i>	<i>2.29 ± 0.00</i>
2500	<i>0.90 ± 0.00</i>	<i>1.46 ± 0.01</i>	<i>2.21 ± 0.03</i>	<i>2.93 ± 0.05</i>	<i>0.88 ± 0.00</i>	<i>1.31 ± 0.01</i>	<i>1.82 ± 0.02</i>	<i>2.20 ± 0.03</i>
3000	<i>0.89 ± 0.00</i>	<i>1.44 ± 0.00</i>	<i>2.14 ± 0.03</i>	<i>2.83 ± 0.04</i>	<i>0.87 ± 0.00</i>	<i>1.29 ± 0.01</i>	<i>1.77 ± 0.02</i>	<i>2.14 ± 0.03</i>
3500	<i>0.89 ± 0.00</i>	<i>1.41 ± 0.01</i>	<i>2.08 ± 0.03</i>	<i>2.76 ± 0.04</i>	<i>0.87 ± 0.00</i>	<i>1.27 ± 0.01</i>	<i>1.72 ± 0.02</i>	<i>2.07 ± 0.02</i>
4000	<i>0.88 ± 0.00</i>	<i>1.40 ± 0.01</i>	<i>2.05 ± 0.02</i>	<i>2.72 ± 0.04</i>	<i>0.86 ± 0.00</i>	<i>1.25 ± 0.01</i>	<i>1.67 ± 0.02</i>	<i>2.00 ± 0.03</i>
4500	<i>0.88 ± 0.00</i>	<i>1.39 ± 0.01</i>	<i>2.00 ± 0.02</i>	<i>2.66 ± 0.04</i>	<i>0.86 ± 0.00</i>	<i>1.24 ± 0.01</i>	<i>1.63 ± 0.02</i>	<i>1.95 ± 0.02</i>
5000	<i>0.88 ± 0.00</i>	<i>1.37 ± 0.01</i>	<i>1.97 ± 0.02</i>	<i>2.62 ± 0.04</i>	<i>0.86 ± 0.00</i>	<i>1.23 ± 0.01</i>	<i>1.60 ± 0.00</i>	<i>1.93 ± 0.02</i>

Results are presented as mean ± SEM, averaged over 50 time series of both Wistar and borderline hypertensive rats, both in baseline and in stress, and averaged over 41 time series of volunteers; results from human volunteers are in italics.

TABLE 5: Level of asymmetry of real signals (skewness).

	Rats		Volunteers	
	SBP	PI	SBP	R-R
Absolute skewness	0.64 ± 0.86	0.48 ± 1.02	0.42 ± 0.54	0.35 ± 0.33
Portion of signals with negative skewness [%]	36%	24%	46%	60%

Absolute skewness is expressed as mean ± σ ; each signal set comprises a couple of highly skewed signals that influence the standard deviation.

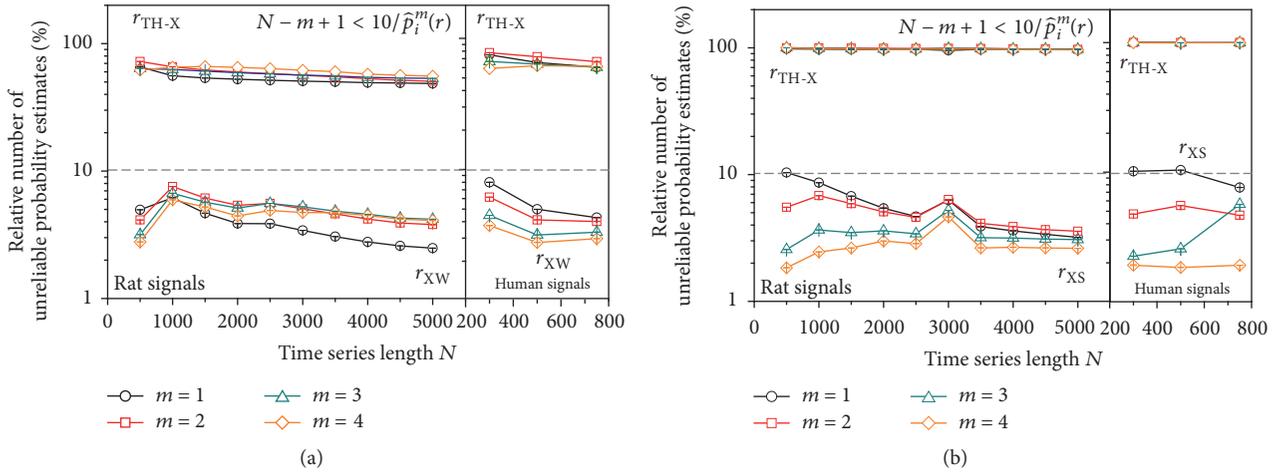


FIGURE 8: Unreliable estimated probabilities $\hat{p}_i^m(r)$ [%] in real signals; (a) $r \in \{r_{XW}, r_{TH-X}\}$ and weak criterion is applied; (b) $r \in \{r_{XS}, r_{TH-X}\}$ and strong criterion is applied. The gray dashed horizontal line marks 10% of the unreliable probability estimates. The left part of each panel shows the results of Wistar and borderline hypertensive rats at baseline and stress; the right part of each panel shows the results of volunteers.

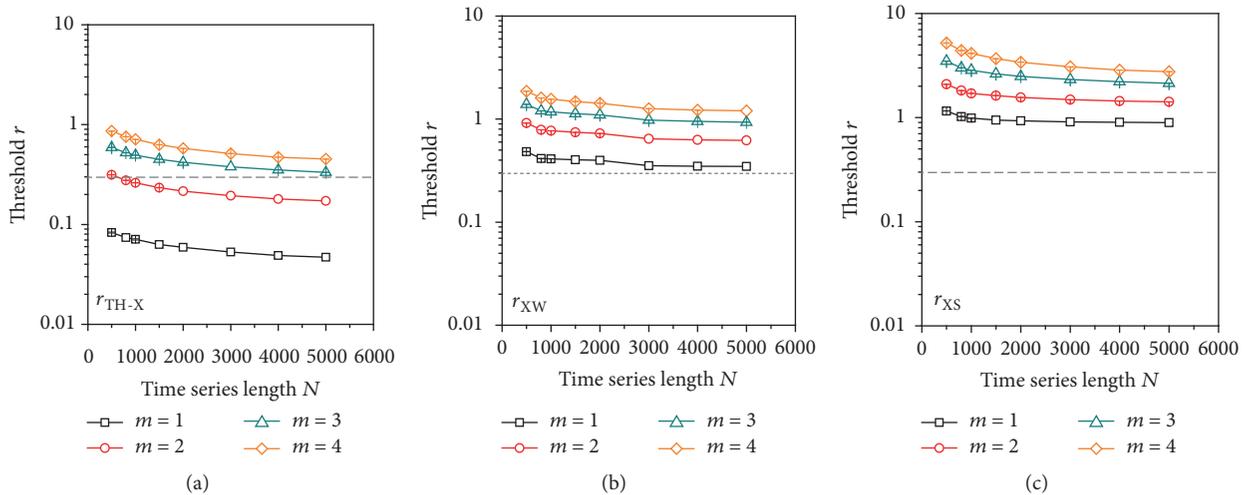


FIGURE 9: Calculated threshold values for artificial data, with signals \mathbf{X} and \mathbf{Y} both normally distributed. (a) Threshold for which $XApEn$ gets maximal value (see (17)). (b) Threshold for which at least 95% of the probabilities satisfy weak criterion (see (19)). (c) Threshold for which at least 95% of the probabilities satisfy strong criterion (see (20)). The gray horizontal dashed line corresponds to $r = 0.3$.

Finally, Table 6 presents a brief recap of the steps necessary for reliable $XApEn$ estimation.

4. Conclusion

This paper performed an analysis of $XApEn$ parameter space and proposed methods for adjusting both the procedure and its parameter choice. Artificial environment with known data distribution allowed exact mathematical evaluation of reference $XApEn$ values and their comparison to the estimated ones. The comparison indicated vast regions in $XApEn$ parameter space (threshold r , data length N , and template length m) with unstable estimation of both $XApEn$ and its constitutive components. It is also shown that parameter interrelationship is nonlinear and depends on data type. An established relationship between the theory

of binary probability estimation [27] and template matching procedure proved that the source of instability is unreliable probability estimation, which is a consequence of working within the unstable region. Stable $XApEn$ estimates should remain constant for increased data lengths and the frequently observed monotonous increase with N is just another form of instability. It is also shown that the traditional guidelines work within the regions of instability. The proposed correction attenuated but did not eliminate the problem.

The solution is to target the parameters within the stable region of the parameter space and to perform the parameter choice jointly, due to the nonlinear properties of their relationship. Real environment does not offer an advantage of exact mathematical error analysis, so a set of parameters for stable and reliable estimation is ensured deriving a novel set of threshold r expressions. The expressions take into

TABLE 6: Steps for reliable $XApEn$ estimation.

Step	Explanation
(1)	Choose two time series and specify which one would be master series; define N and m (preferably $m = 2$); perform standard scoring of both series.
(2)	Find the differential time series both for master and for follower time series, and calculate their respective standard deviations σ_{DX} and σ_{DY} .
(3)	Calculate r_{TH-A} (see (18)).
(4)	Calculate r_{TH-X} (see (17)).
(5)	Select reliability criterion; a strong criterion requires long data sets, so a weak criterion is suggested.
(6)	Calculate r_{XW} (see (19)), weak criterion (preferably); or calculate r_{XS} (see (20)), strong criterion.
(7)	Estimate $XApEn(N, m, r_{XW})$.

account length N , template m , statistical data properties, and reliability criterion (weak or strong, [27]). The consistency of estimates is validated for typical cases of short records (ambulatory patients) and long records (small laboratory animals). It is shown that classical guidelines underestimate the threshold levels and data lengths required for consistent results.

This analysis can be employed for all the entropy measures that rely on probability estimation. The proposed $XApEn$ improvements may contribute to its wider implementation in the domain of related physiological processes such as baroreflex studies. In order to facilitate and promote the wider $XApEn$ deployment, we have included the detailed instructions within the Supplementary Material.

Although the method is validated in cardiovascular data, it can be equivalently applied for finding correct parameters for $XApEn$ analysis of processes that are not necessarily biomedical, provided that their distribution is not too skewed. Further investigation would consider the signals with highly asymmetric distribution, for example, interspike interval time series, but also the signals with mismatched distribution where zero probability is not an error, but inherent signal property.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

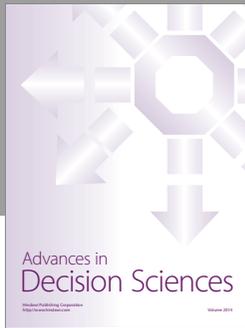
Acknowledgments

This work was supported by the Serbian Ministry of Education, Science and Technology Development, under Grants TR32040 and III41013. The authors acknowledge the support from the EU COST-Action CA15104-Body Communications Group. The authors are also grateful to Professor Dr. Harald Straus, University of Iowa, and to Professor Dr. Suave Lobodzinski, UCLA, for the inspiring discussions in October 2016.

References

- [1] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 6, pp. 2297–2301, 1991.
- [2] S. Pincus, "Approximate entropy (ApEn) as a complexity measure," *Chaos*, vol. 5, no. 1, pp. 110–117, 1995.
- [3] S. Pincus and B. H. Singer, "Randomness and degrees of irregularity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 5, pp. 2083–2088, 1996.
- [4] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, and N. Stergiou, "The appropriate use of approximate entropy and sample entropy with short data sets," *Annals of Biomedical Engineering*, vol. 41, no. 2, pp. 349–365, 2013.
- [5] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology—Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.
- [6] A. Porta, G. Baselli, D. Liberati et al., "Measuring regularity by means of a corrected conditional entropy in sympathetic outflow," *Biological Cybernetics*, vol. 78, no. 1, pp. 71–78, 1998.
- [7] A. Porta, P. Castiglioni, V. Bari et al., "K-nearest-neighbor conditional entropy approach for the assessment of the short-term complexity of cardiovascular control," *Physiological Measurement*, vol. 34, no. 1, article no. 17, pp. 17–33, 2013.
- [8] W. T. Chen, Z. Z. Wang, H. B. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 2, pp. 266–272, 2007.
- [9] M. Costa, A. L. Goldberger, and C. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E*, vol. 71, no. 2, Article ID 021906, 2005.
- [10] P. Li, C. Liu, K. Li, D. Zheng, C. Liu, and Y. Hou, "Assessing the complexity of short-term heartbeat interval series by distribution entropy," *Medical and Biological Engineering and Computing*, vol. 53, no. 1, pp. 77–87, 2015.
- [11] T. Skoric, O. Mohamoud, B. Milovanovic, N. Japundzic-Zigon, and D. Bajic, "Binarized cross-approximate entropy in crowd-sensing environment," *Computers in Biology and Medicine*, vol. 80, pp. 137–147, 2017.
- [12] X. Ning, Y. Xu, J. Wang, and X. Ma, "Approximate entropy analysis of short-term HFECG based on wave mode," *Physica A: Statistical Mechanics and its Applications*, vol. 346, no. 3–4, pp. 475–483, 2005.
- [13] Y.-H. Pan, Y.-H. Wang, S.-F. Liang, and K.-T. Lee, "Fast computation of sample entropy and approximate entropy in biomedicine," *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, pp. 382–396, 2011.
- [14] S. Zurek, P. Guzik, S. Pawlak, M. Kosmider, and J. Piskorski, "On the relation between correlation dimension, approximate entropy and sample entropy parameters, and a fast algorithm for their calculation," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 24, pp. 6601–6610, 2012.
- [15] S. Lu, X. Chen, J. K. Kanters, I. C. Solomon, and K. H. Chon, "Automatic selection of the threshold value r for approximate entropy," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, article no. 8, pp. 1966–1972, 2008.
- [16] K. Chon, C. Scully, and S. Lu, "Approximate entropy for all signals," *IEEE Engineering in Medicine and Biology Magazine*, vol. 28, no. 6, pp. 18–23, 2009.
- [17] P. Castiglioni and M. Di Rienzo, "How the threshold ' r ' influences approximate entropy analysis of heart-rate variability," in

- Proceedings of the Computers in Cardiology 2008, CAR*, pp. 561–564, Bologna, Italy, September 2008.
- [18] C. Liu, C. Liu, P. Shao et al., “Comparison of different threshold values r for approximate entropy: Application to investigate the heart rate variability between heart failure and healthy control groups,” *Physiological Measurement*, vol. 32, no. 2, pp. 167–180, 2011.
- [19] J. F. Restrepo, G. Schlotthauer, and M. E. Torres, “Maximum approximate entropy and r threshold: A new approach for regularity changes detection,” *Physica A: Statistical Mechanics and its Applications*, vol. 409, pp. 97–109, 2014.
- [20] A. Boskovic, T. Loncar-Turukalo, O. Sarenac, N. Japundzic-Zigon, and D. Bajic, “Unbiased entropy estimates in stress: A parameter study,” *Computers in Biology and Medicine*, vol. 42, no. 6, pp. 667–679, 2012.
- [21] F. Kaffashi, R. Foglyano, C. G. Wilson, and K. A. Loparo, “The effect of time delay on approximate & sample entropy calculations,” *Physica D. Nonlinear Phenomena*, vol. 237, no. 23, pp. 3069–3074, 2008.
- [22] S. M. Pincus, T. Mulligan, A. Iranmanesh, S. Gheorghiu, M. Godschalk, and J. D. Veldhuis, “Older males secrete luteinizing hormone and testosterone more irregularly, and jointly more asynchronously, than younger males,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 24, pp. 14100–14105, 1996.
- [23] P. Li, K. Li, C. Liu, D. Zheng, Z.-M. Li, and C. Liu, “Detection of coupling in short physiological series by a Joint Distribution Entropy Method,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 11, pp. 2231–2242, 2016.
- [24] O. Šarenac, M. Lozić, S. Drakulić et al., “Autonomic mechanisms underpinning the stress response in borderline hypertensive rats,” *Experimental Physiology*, vol. 96, no. 6, pp. 574–589, 2011.
- [25] T. Ceranic, T. Loncar-Turukalo, B. Milovanovic, N. Japundzic-Zigon, and D. Bajic, “Cross-entropy of systolic blood pressure - Pulse interval: Automatic threshold and its reliability,” in *Proceedings of the 2013 40th Computing in Cardiology Conference, CinC 2013*, pp. 1219–1222, Zaragoza, Spain, September 2013.
- [26] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen, “An advanced detrending method with application to HRV analysis,” *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 2, pp. 172–175, 2002.
- [27] M. C. Jeruchim, “Techniques for Estimating the Bit Error Rate in the Simulation of Digital Communication Systems,” *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 1, pp. 153–170, 1984.
- [28] H.-T. Wu, C.-Y. Lee, C.-C. Liu, and A.-B. Liu, “Multiscale cross-approximate entropy analysis as a measurement of complexity between ECG R-R interval and PPG pulse amplitude series among the normal and diabetic subjects,” *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 231762, 2013.
- [29] S. Alesci, P. E. Martinez, S. Kelkar et al., “Major depression is associated with significant diurnal elevations in plasma interleukin-6 levels, a shift of its circadian rhythm, and loss of physiological complexity in its secretion: Clinical implications,” *Journal of Clinical Endocrinology and Metabolism*, vol. 90, no. 5, pp. 2522–2530, 2005.
- [30] H.-T. Wu, C.-C. Liu, M.-T. Lo et al., “Multiscale cross-approximate entropy analysis as a measure of complexity among the aged and diabetic,” *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 324325, 7 pages, 2013.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

