Characterization of Physiological Noise in Complex Cardiovascular Variability Series

Andrea Scarciglia[1], Vincenzo Catrambone[1], Claudio Bonanno[2], and Gaetano Valenza[1]

***Abstract*— The cardiovascular system can be analyzed using spectral, nonlinear, and complexity metrics. Nevertheless, dynamical noise may significantly impact these quantifiers. To our knowledge, there has been no attempt to quantify the intrinsic cardiovascular system noise driving heartbeat dynamics. To this end, this study presents a novel, model-free framework to define and quantify physiological noise using nonlinear Approximate Entropy profile. The framework was tested using analytical noisy series and then applied to real Heart Rate Variability (HRV) series gathered from a publicly-available dataset of recordings from 19 young and 19 elderly subjects watching the movie "Fantasia". Results suggest that physiological noise may account for over 15% of cardiovascular dynamics and is influenced by aging, with decreased cardiac noise in the elderly compared to the young subjects. Our findings indicate that physiological noise is a crucial factor in characterizing cardiovascular dynamics, and current spectral, nonlinear, and complexity assessments should take into account underlying dynamical noise estimates.**

I. Introduction

The analysis of Heart Rate Variability (HRV) series provides valuable information on the autonomic regulation of cardiovascular dynamics and has been used to study various pathological states, including mood disorders, cardiovascular diseases, and aging [1]–[5]. Accordingly, nonlinear and complexity measures, such as the Lyapunov exponent [6], embedding dimensions [7], and entropies [3], [8], are increasingly popular for HRV analysis [3].

Despite the ability of nonlinear and complexity analyses to differentiate these pathological and physiological conditions, their application in clinical settings is limited by the absence of clear physiological correlates and specificity issues [9], [10]. HRV series represent the time intervals between consecutive heartbeats, as indicated by the R peaks in the electrocardiogram (ECG), and are thought to be the output of a nonlinear system showing complex dynamics, being influenced by intrinsic dynamical noise [1], [2], [11]. Such intrinsic physiological noise may be due to the interaction between the autonomic nervous system sub-components, alongside with other physiological systems (e.g., respiration, central nervous system) [12], [13], and may significantly alter complexity estimates and related pathophysiological interpretations [14], [15]. While the role of physiological noise in biological processes has been acknowledged [16], [17], to the best of our knowledge, it has not been formally defined nor has cardiovascular noise been calculated in HRV series.

In this study, we aim to formally define in a closed form and estimate physiological noise in the cardiovascular system through HRV series analysis. Assuming HRV series result from a dynamical combination of deterministic and stochastic components, we model physiological noise as independent and identically (IID) distributed random variables with a Gaussian distribution $\mathcal{N}(0,\sigma^2)$ and estimate the standard deviation σ by exploiting the approximate entropy (ApEn) [18] profile. Our model-free method for dynamical noise estimation was chosen over well-known model-based techniques, such as the Kalman filter or autoregressive models, as it does not make any assumptions about the underlying (deterministic) system dynamics [19]. We first demonstrate the accuracy and robustness of the proposed noise estimation method on synthetic time series with varying levels of dynamical noise, simulating the effects of physiological noise. We then estimate physiological noise in HRV series obtained from young and elderly subjects.

II. Materials and Methods

A. Physiological Noise Definition and Physiological Noise Estimation Procedure

Let us represent the cardiovascular system as a discrete metrical dynamical system (X,μ,T), which is defined by any piecewise differentiable map function T with bounded derivative on a compact set $X \subset \mathbb{R}$ and preserving the probability measure μ. A generic noise-free orbit of such a dynamical system is of the form $w_n = T(w_{n-1}, w_{n-2}, ..., w_0)$, with $w_i \in X$, for all positive integers i.

We define the *physiological noise* as the *dynamical noise* constituted by a sequence of IID Gaussian random variables $\{\varepsilon_n\}_n$ following the distribution $\mathcal{N}(0,\sigma^2)$, and whose samples modify physiological dynamics at each step according to the following equation:

$$x_n = T(x_{n-1}, x_{n-2}, ..., x_0) + \varepsilon_n, \tag{1}$$

with $x_i \in X$ for all positive integers i, and $\{x_n(\sigma)\}_{n=1}^N$ a noisy physiological time series comprising N samples (next, a HRV series).

According to the embedding theorem [20], the reconstructed m-dimensional state space $\mathscr{M}_x$ of X is as follows:

$$\mathscr{M}_x(j) = \{x(j), x(j+\tau), ..., x(j+(m-1)\tau)\} \tag{2}$$

[1]Bioengineering and Robotics Research Center E.Piaggio & Department of Information Engineering, School of Engineering, University of Pisa, Italy;
[2]Department of Mathematics, University of Pisa, Italy;
*Corresponding author. Email at: *andrea.scarciglia@phd.unipi.it*

The research leading to these results has received partial funding from the European Commission under grant agreement N. 101017727 for the project EXPERIENCE and by Italian Ministry of Education and Research (MIUR) in the framework of the FoReLab project (Departments of Excellence).

where $j \in 1,2,...,N$ is a time index, m is the embedding dimension, and τ is the delay-time for the state-space reconstruction, and $N+(m-1)\tau$ is time series cardinality.

In accordance with the theory fully reported in [19], we here derive the standard deviation σ of the physiological noise in a closed form as follows:

$$\log(\sigma) \approx \mathrm{ApEn}(X,m,r) + \log(r\sqrt{\pi})$$

A formal definition of the well-known ApEn [18] is reported in the Appendix I below. Assuming that a physiological time series $\{x_n(\sigma)\}_{n=1}^{N}$ under investigation satisfies Equation 1, meaning it is the output of an unknown discrete metric dynamical system contaminated by dynamical noise, the standard deviation of the physiological noise (σ) can be estimated through the following steps:

- A fixed embedding dimension m is chosen;
- The map $r \to \mathrm{ApEn}(\{x_n(\sigma)\}_{n=1}^{N}, m, r)$ is computed as a function of the tolerance parameter r, which varies from 0 to the amplitude of series X with a step Δr of arbitrary size;
- A rough estimate of σ is obtained by identifying $\bar{r}$, the value that minimizes the discrete derivative of $r \to \mathrm{ApEn}(\{x_n(\sigma)\}_{n=1}^{N}, m, r) + \log r$;
- Within the interval $I(\bar{r}) = [r_{max}, \bar{r}]$, where r_{max} is the tolerance value at which $r \to \mathrm{ApEn}(\{x_n(\sigma)\}_{n=1}^{N}, m, r)$ reaches its maximum, the best fit $\bar{\sigma}$ for the function $\sigma \to \mathrm{ApEn}(\{x_n(\sigma)\}_{n=1}^{N}, m, r) + \log[r/(\sigma\sqrt{\pi})]$ is determined.

In this paper, all noise estimates are performed with a resolution step of $\Delta r = \{\text{series amplitude}\}/1000$.

B. Synthetic Data

The proposed estimation method is here evaluated using analytical series generated from the Logistic map, defined as

$$f_\lambda : [0,1] \to [0,1], \quad f_\lambda(x) = \lambda x(1-x)$$

where λ is a real-valued parameter in the interval $[0,4]$ and set to 3.5 in periodic regime. To simulate the effect of noise, 100 perturbed series with various standard deviation percentages ($2\%, 5\%, 10\%, 15\%, 20\%$) relative to the amplitude of the noise-free series ($N = 5000$ samples) were generated. The noise-free series were created using a random initial condition $x_0 \in [0,1]$. A white Gaussian process series (5000 samples) was then added to the noise-free series with the standard deviation as the specified percentage of the amplitude. The perturbed series were constructed by adding a sample of the noise realization at each step of the Logistic map, taking into account the dynamical noise law and applying reduction modulo 1 at each step to ensure that the points of the map lie in the interval $[0,1]$. This process is referred to as the "bounce" effect.

C. Experimental Setup and Data Preprocessing

This study was approved by the committee of bioethics of the University of Pisa with review n. 19/2021. A total of 38 ECG signals were obtained from two groups of healthy subjects: 20 young adults aged 21 to 34 years old and 20 elderly adults aged 68 to 85 years old. The ECG signals were collected from the publicly available *Fantasia* database [21] and are accessible at *https://physionet.org/content/fantasia/1.0.0/*. The two groups contained equal numbers of males and females. The continuous ECG signals were recorded at a frequency of 250Hz during 120 minutes of rest, with subjects required to remain awake by watching the movie *Fantasia* (Disney, 1940) [21]. The HRV series were derived from the ECG signals using the Pan-Tompkins algorithm [22] and were verified through visual inspection. The cardinality N in HRV series of the young subjects varies from N=5358 to N=9053 (median: N=7100), while that of the elderly cohort ranges from N=6752 to N=8410 (median: N=6752).

D. Statistical Analysis

The comparison of standard deviation noise estimates between the 5 different levels of noise (2%, 5%, 10%, 15%, 20%) in the synthetic series was conducted using a non-parametric Kruskal-Wallis test. The null hypothesis was equal median among populations, with a significance level of $\alpha = 0.01$. Additionally, multiple pairwise comparisons were performed using Mann-Whitney tests for unpaired samples, with a significance level of $\alpha_1 = \alpha/10 = 0.001$.

For the artifact-free HRV series in the dataset "Fantasia", non-parametric Mann-Whitney tests for unpaired samples were used to compare the standard deviation noise estimates between the elderly and young groups. The null hypothesis was equal median between the two populations. The statistical comparison was performed for both embedding dimensions $m = 2$ and $m = 3$, with a significance level of $\alpha = 0.01$.

III. Experimental Results

The results of the noise estimation method applied to the analytical logistic series benchmark are displayed in Fig. 1 in the form of boxplot statistics. The noise standard deviation was estimated for an embedding dimension of $m = 2$. The estimated values of the dynamical noise are shown with respect to the actual, superimposed noise for each of the 5 groups with different levels of standard deviation.

A visual examination of the results suggests that the algorithm accurately detects and estimates the different levels of superimposed dynamical noise across all realizations, as indicated by the small dispersion around the median values. It is noted that as the level of superimposed noise decreases, the estimated noise values are more consistently concentrated around the actual value. Although noisier series exhibit a slight increase in dispersion, it is possible that this is due to the similarity of the corrupted dynamics with a pure stochastic process. However, this increase in dispersion does not significantly affect the accuracy of the noise estimation, as evidenced by the p-values obtained from the Kruskal-Wallis and multiple pairwise Mann-Whitney tests for unpaired groups with a null hypothesis of equal median among populations. These p-values were found to be less than 10^{-6}.

The results of the noise estimation method on physiological data are presented as boxplots in Fig. 2 for embedding

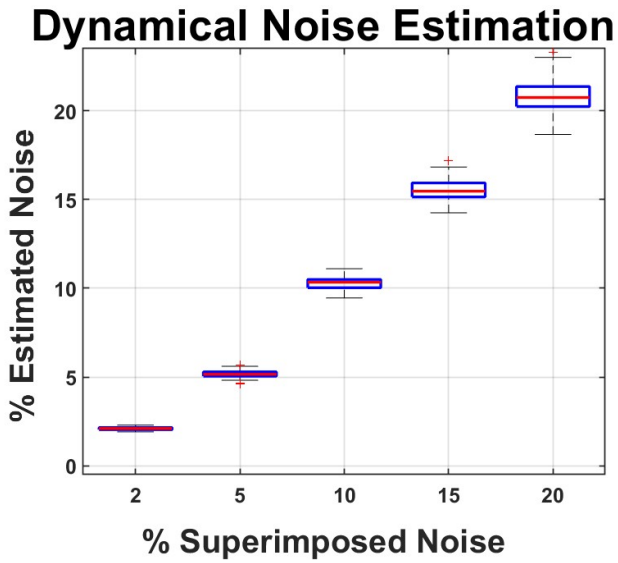

Fig. 1: Dynamical noise standard deviation estimation at different levels for Logistic map, λ=3.5. Noise is estimated at fixed embedding dimension ($m = 2$). Series length is of 5000 samples for 100 realizations. The percentages of imposed and estimated noise correspond to the ratio between the noise std value and the noise-free signal amplitude. P-values of Kruskal-Wallis and multiple pairwise Mann-Whitney tests for unpaired group with a null-hypothesis of equal median among populations are lower than 10^{-6}.

dimensions $m = 2$ and $m = 3$. The estimates of physiological noise do not exhibit significant variations with respect to the embedding dimension. A comparison of the absolute noise standard deviation achieved with the proposed method for the elderly and young groups is shown in the upper boxes of the figure. In terms of median $\pm$ median absolute deviation (MAD), the results indicate that young people have significantly higher cardiac noise levels (0.0454 ± 0.0188 for $m = 2$ and 0.0453 ± 0.0153 for $m = 3$) compared to elderly subjects (0.0219 ± 0.0058 for $m = 2$ and 0.0208 ± 0.0054 for $m = 3$), with $p < 0.01$. This difference was expected as elderly subjects tend to have a lower range in their HRV series, leading to a reduced noise level.

When the absolute noise estimation is subject-wise standardized with respect to the HRV series amplitude (as if all series have an amplitude equal to 1), the results show that young people still have higher noise standard deviation ($10.45\% \pm 2\%$ for $m = 2$ and $9.53\% \pm 1.6\%$ for $m = 3$) compared to elderly subjects ($7.2\% \pm 2\%$ for $m = 2$ and $7.15\% \pm 1.5\%$ for $m = 3$), regardless of the embedding dimension. Furthermore, noise levels in young groups exhibit more variability, as evidenced by the larger dispersion around the median value.

IV. Discussion and Conclusion

We provide a methodological framework to formally define and estimate physiological noise in cardiovascular systems. HRV series are assumed to be the outputs of an unknown dynamical system that results from the interplay between deterministic and stochastic components. While we do not make any assumptions regarding the deterministic cardiovascular dynamics, we assume that the stochastic component - physiological noise - can be modeled as a sequence of IID random variables following a Gaussian distribution $\mathcal{N}(0, \sigma^2)$. This modeling choice is justified by the fact that physiological noise is expected to affect all the frequencies of the HRV spectrum equally, rather than being concentrated in a specific frequency band (as in the case of colored noise). Consequently, the proposed dynamical noise estimation algorithm is suitable for HRV series as it does not require any specific assumptions about the underlying

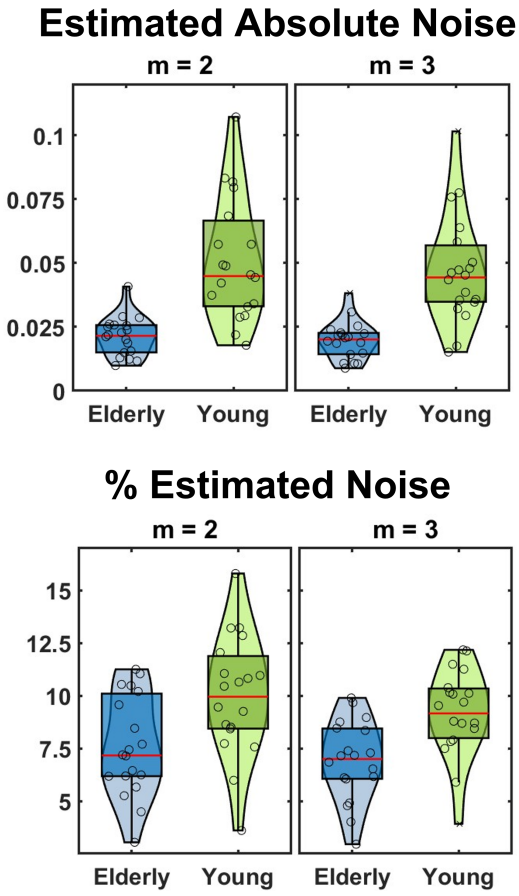

Fig. 2: Noise estimation in cardiovascular variability series. Boxplots statistics for the proposed physiological noise estimation in Elderly vs Young groups: absolute noise (top figures) and standardized noise with respect to the amplitude of the series (figures below). Estimates used embedding dimension $m = 2$ and $m = 3$. Red segments stand for median values, circles denote the single-subject estimate, crosses the outlier value points, while "violin" plot denotes the estimate's density across central values. For both absolute and standardized noise, young cohort subjects shows a higher median value with respect to the elderly ones ($p < 0.01$ for both embedding dimensions).

dynamics, which may be constituted by unknown nonlinear functions showing complex dynamics [1]–[5], [11], [12].

Experimentally, we first demonstrate the effectiveness of the method on analytical benchmarks, consisting of synthetic series generated by the logistic map in periodic regime and corrupted by various levels of dynamical noise. Indeed, both logistic maps and HRV series are output of discrete-time, nonlinear systems that generate time-series data. Our numerical results confirmed the accuracy of the method in distinguishing the different levels of noise.

We then apply the proposed noise estimation method to 38 HRV series from the publicly available Fantasia database, including 19 series from young and 19 from elderly subjects in a resting state. Previous studies have suggested that complexity in heartbeat dynamics decreases with age [23], [24], but these measures are biased by the assumption of a noise-free cardiovascular system dynamics. Our numerical analysis shows that aging is associated with lower noise levels in autonomic control on heartbeat dynamics, as the noise standard deviation decreases in the elderly group with less inter-subject variability. This phenomenon suggests that aging may regulate physiological noise rather than directly modulating the underlying system dynamics. These results are consistent with previous studies [25] where aging was found to have no effect on complex cardiovascular dynamics. Accordingly, we may hypothesize that the presence of physiological noise, which was not taken into account in previous complexity assessment, has strongly altered the results in the literature regarding complex cardiovascular dynamics [1]–[4], [14]. This study has some limitations, including the small sample size and the need for further investigation with larger cohorts and in different physiological and pathological conditions. Moreover, the physiological mechanisms responsible for the observed age-related changes in noise levels of cardiovascular dynamics require further investigation. Note that the proposed physiological noise estimation may be applied to any time series generated by a complex dynamical system, therefore it may be applied to calculate noise in physiological systems other than the cardiovascular one, e.g., brain/cortical/neural noise.

V. Appendix I: Approximate Entropy (ApEn)

The nonlinear quantifier Approximate Entropy (ApEn) [18] is non-negative and quantifies the predictability of a time series. Formally, given a series $\{y_n\}_{n=1}^N$ of N samples and a positive integer m, the series is embedded in $\mathbb{R}^m$ forming vectors $Y_i = (y_i, \ldots, y_{i+m-1})$ with $i = 1, \ldots, N-m+1$. For a distance d in $\mathbb{R}^m$ and for a positive value r, we define $C_i^m(r) = \{\text{number of } j \text{ s.t. } d(Y_j, Y_i) < r\}$ and $\Phi^m(r) = (N-m+1)^{-1} \sum_i^{N-m+1} \log C_i^m(r)$, which yield the definition of the ApEn:

$$\text{ApEn}(\{y_n\}_{n=1}^N, m, r) = \Phi^m(r) - \Phi^{m+1}(r).$$

In this setting, the $\text{ApEn}(\{y_n\}_{n=1}^N, m, r)$ assesses the proximity of the embedded vectors $\{Y_j\}_{j=1}^{N-m+1}$.

References

[1] U. R. Acharya *et al.*, "Heart rate variability: a review," *Medical and Biological Engineering and Computing*, vol. 44, no. 12, pp. 1031–1051, 2006.

[2] R. Barbieri *et al.*, *Complexity and nonlinearity in cardiovascular signals*. Springer, 2017.

[3] R. Sassi *et al.*, "Advances in heart rate variability signal analysis: joint position statement by the e-cardiology esc working group and the european heart rhythm association co-endorsed by the asia pacific heart rhythm society," *Ep Europace*, vol. 17, no. 9, pp. 1341–1353, 2015.

[4] A. L. Goldberger *et al.*, "What is physiologic complexity and how does it change with aging and disease?," *Neurobiol Aging*, vol. 23, no. 1, pp. 23–26, 2002.

[5] L. Glass, "Introduction to controversial topics in nonlinear science: Is the normal heart rate chaotic?," 2009.

[6] G. Valenza, L. Citi, and R. Barbieri, "Estimation of instantaneous complex dynamics through lyapunov exponents: a study on heartbeat dynamics," *PloS one*, vol. 9, no. 8, p. e105622, 2014.

[7] R. Bhavsar *et al.*, "Time series analysis using embedding dimension on heart rate variability," *Procedia computer science*, vol. 145, pp. 89–96, 2018.

[8] M. Costa *et al.*, "Multiscale entropy analysis of complex physiologic time series," *Physical review letters*, vol. 89, no. 6, p. 068102, 2002.

[9] F. Beckers *et al.*, "Effects of autonomic blockade on non-linear cardiovascular variability indices in rats," *Clinical and experimental pharmacology and physiology*, vol. 33, no. 5-6, pp. 431–439, 2006.

[10] J. P. Saul and G. Valenza, "Heart rate variability and the dawn of complex physiological signal analysis: methodological and clinical perspectives," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2212, p. 20200255, 2021.

[11] A. A. Armoundas *et al.*, "A stochastic nonlinear autoregressive algorithm reflects nonlinear dynamics of heart-rate fluctuations," *Annals of biomedical engineering*, vol. 30, no. 2, p. 192, 2002.

[12] K. Sunagawa *et al.*, "Dynamic nonlinear vago-sympathetic interaction in regulating heart rate," *Heart and vessels*, vol. 13, pp. 157–174, 1998.

[13] V. Catrambone *et al.*, "Functional brain–heart interplay extends to the multifractal domain," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2212, p. 20200260, 2021.

[14] C. K. Rhea *et al.*, "Noise and complexity in human postural control: interpreting the different estimations of entropy," *PloS one*, vol. 6, no. 3, p. e17696, 2011.

[15] D. Chelidze, "Reliable estimation of minimum embedding dimension through statistical analysis of nearest neighbors," *Journal of Computational and Nonlinear Dynamics*, vol. 12, no. 5, 2017.

[16] A. A. Faisal *et al.*, "Noise in the nervous system," *Nature reviews neuroscience*, vol. 9, no. 4, pp. 292–303, 2008.

[17] E. Sejdić *et al.*, "Necessity of noise in physiology and medicine," *Computer methods and programs in biomedicine*, vol. 111, no. 2, pp. 459–470, 2013.

[18] S. Pincus, "Approximate entropy as a measure of system complexity.," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[19] A. Scarciglia *et al.*, "Estimation of dynamical noise power in unknown systems," *IEEE Signal Processing Letters*, 2023.

[20] L. Noakes, "The takens embedding theorem," *International Journal of Bifurcation and Chaos*, vol. 1, no. 04, pp. 867–872, 1991.

[21] N. Iyengar *et al.*, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 271, no. 4, pp. R1078–R1084, 1996.

[22] J. Pan and W. Tompkins, "A real-time qrs detection algorithm," *IEEE transactions on biomedical engineering*, no. 3, pp. 230–236, 1985.

[23] D. M. Simpson and R. Wicks, "Spectral analysis of heart rate indicates reduced baroreceptor-related heart rate variability in elderly persons," *Journal of Gerontology*, vol. 43, no. 1, pp. M21–M24, 1988.

[24] H. Tsuji *et al.*, "Reduced heart rate variability and mortality risk in an elderly cohort. the framingham heart study.," *Circulation*, vol. 90, no. 2, pp. 878–883, 1994.

[25] D. T. Schmitt *et al.*, "Fractal scale-invariant and nonlinear properties of cardiac dynamics remain stable with advanced age: a new mechanistic picture of cardiac control in healthy elderly," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 293, no. 5, pp. R1923–R1937, 2007.