# Instantaneous Nonlinear Assessment of Complex Cardiovascular Dynamics by Laguerre-Volterra Point Process Models

Gaetano Valenza, Member, IEEE, Luca Citi<sup>\*</sup>, Member, IEEE, and Riccardo Barbieri, Senior Member, IEEE

*Abstract*— We report an exemplary study of instantaneous assessment of cardiovascular dynamics performed using pointprocess nonlinear models based on Laguerre expansion of the linear and nonlinear Wiener-Volterra kernels. As quantifiers, instantaneous measures such as high order spectral features and Lyapunov exponents can be estimated from a quadratic and cubic autoregressive formulation of the model first order moment, respectively. Here, these measures are evaluated on heartbeat series coming from 16 healthy subjects and 14 patients with Congestive Hearth Failure (CHF). Data were gathered from the on-line repository PhysioBank, which has been taken as landmark for testing nonlinear indices. Results show that the proposed nonlinear Laguerre-Volterra pointprocess methods are able to track the nonlinear and complex cardiovascular dynamics, distinguishing significantly between CHF and healthy heartbeat series.

#### I. INTRODUCTION

High Order Spectra (HOS) and Lyapunov Exponents (LEs) are widely recognized as important quantifiers of, respectively, nonlinearity and complexity. HOS are spectral representations of higher order statistics, i.e. moments and cumulants of third order and beyond [1]. HOS can detect deviations from linearity, stationarity or Gaussianity in the signal. A particular case is represented by the third-order spectrum, also called bispectrum, which is by definition the Fourier transform of the third-order cumulant sequence. LEs are nonlinear measures useful for the characterization of complexity in a nonlinear system. In such a case, complex dynamics refer to the possible chaotic behavior of the system variables. In a stable purely deterministic nonlinear system, for instance, positive LEs reflect a strong dependence to initial conditions leading to the definition of chaotic system.

Both HOS and LEs have been successfully estimated on time series coming from physiological systems, in which nonlinear behavior plays a crucial role [2]. In particular, these concepts have been applied on the series mostly used for autonomic and cardiovascular assessment, i.e., heart rate variability (HRV) [2]. HRV is the analysis of the variability of the RR series, the distance between two heartbeat events identified as R-waves extracted from the electrocardiogram. HOS and LEs estimation on HRV is not a trivial task. Most of the proposed methods, in fact, require interpolation of unevenly-spaced samples, introducing algorithmic artifacts.

G. Valenza, L. Citi, and R. Barbieri are with the Neuroscience Statistics Research Laboratory, Harvard Medical School, Massachusetts General Hospital, Boston, MA, 02114 USA, and also with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: g.valenza@ieee.org; {lciti,barbieri}@neurostat.mit.edu).

L. Citi and G. Valenza equally contributed to this work.

Moreover, to obtain reliable estimations, relatively long-time observations are needed.

To overcome these issues, a point process-based approach on heartbeat series has been recently proposed [3]–[5] allowing for instantaneous cardiovascular and autonomic assessment without any interpolation requirement. The pointprocess model's effectiveness has been validated against other methods, both for linear [3] and nonlinear [6] formulations. This stochastic approach defines the probability of having a heartbeat event at each moment in time. Accordingly, each RR interval is characterized by an inverse-gaussian (IG) distribution whose first order moment is defined with a fully parametric and autoregressive formulation. Alongside the simple autoregressive linear combination of the present and past RR intervals, Nonlinear Autoregressive (NAR) terms can also be taken into account in the formulation of the IG mean. Specifically, previous works linked the instantaneous HOS estimations to second-order nonlinearities [4] and instantaneous LEs to third-order nonlinearities [5] defining the so-called instantaneous dominant Lyapunov exponent (IDLE). Due to the intrinsic long-term memory of the cardiovascular system and the high order of nonlinearities, a large number of parameters needs to be estimated. Therefore, the linear and nonlinear autoregressive kernels, formulated as Wiener-Volterra series, are expanded using the discretetime Laguerre bases [7], devising the quadratic and cubic nonlinear autoregressive Laguerre models (hereinafter  $2^{nd}$ NARL and  $3^{rd}$  NARL, respectively).

In this work we report both NAR and NARL applications considering HOS and IDLE of heartbeat series gathered from healthy subjects and patients with Congestive Heart Failure (CHF). Such series belong to the well-known online database *physiobank* [8]. The primary goal is to study the role of the mentioned instantaneously derived features in discriminating the control and CHF groups with statistical significance. The assessment of HRV chaotic behavior is not addressed here. It has been demonstrated, in fact, that LEs estimations on RR interval series are not reliable in terms of chaotic assessment, mainly because of the presence of intrinsic stochastic terms in the cardiovascular control dynamics [9]. Therefore, according to current literature, claims on negative LE values are associated with predictable and less complex cardiovascular dynamics, whereas positive LE values indicate more complex and unpredictable dynamics.

## II. METHODS

We report essential methodological details on the models involved in this study. Specifically, starting from the general

<sup>\*</sup> Corresponding author.

This research was supported in part by NIH Grant R01-HL084502.

point process formulation for human heartbeat (see details in [3]), the  $2^{nd}$  NARL and NAR derivation (see details in [4]) is seen as further simplification of the  $3^{rd}$  NARL model (see details in [5]). Then, the estimation of the IDLE, using the  $3^{rd}$  NARL model, and the instantaneous bispectral features, using the  $2^{nd}$  NARL or NAR model, are described along.

Considering an ECG observation in the time domain, within the interval  $t \in (0, T]$ , it is possible to define  $RR_{\tilde{N}(t)}$  as the j<sup>th</sup> R–R interval, with  $j = \tilde{N}(t)$  the index of the previous R-wave event before time  $t$  related to the cardiac left-continuous counting process  $N(t)$ . Assuming history dependence, a physiologically-plausible IG distribution,  $f(t|\mathcal{H}_t, \xi(t))$ , can be adopted to characterize the probability distribution of the waiting time from  $t$  to the next R-wave event [3].  $\mathcal{H}_t$  stands for the mathematical representation of the past R-events, and  $\xi(t)$  represents the vector of the model parameters including the shape parameter of the inverse Gaussian distribution. Rather than defining the IG mean  $\mu_{\text{RR}}(t, \mathcal{H}_t, \xi(t))$  as direct nonlinear regression of the previous RR intervals, the Laguerre functions [7] are used here to expand the kernels and reduce the number of unknown parameters that need be estimated. The generic  $i^{\text{th}}$ order discrete time Laguerre function,  $\phi_i(n)$  [7] is defined as:

$$
\phi_i(n) = \alpha^{\frac{n-i}{2}} (1 - \alpha)^{\frac{1}{2}} \sum_{j=0}^i (-1)^j {k \choose j} {i \choose j} \alpha^{i-j} (1 - \alpha)^j
$$

with  $(n \geq 0)$  and  $\alpha \in (0,1)$  the discrete-time Laguerre parameter which determines the rate of exponential asymptotic decline of these functions. The instantaneous  $\mu_{RR}(t, \mathcal{H}_t, \xi(t))$  for the  $3^{rd}$  NARL model becomes:

$$
\mu_{RR}(t, \mathcal{H}_t, \xi(t)) = g_0(t) + \sum_{i=0}^p g_1(i, t) l_i(t) + \sum_{i=0}^q \sum_{j=0}^q g_2(i, j, t) l_i(t) l_j(t) + \sum_{i=0}^k \sum_{j=0}^k \sum_{k=0}^k g_3(i, j, k, t) l_i(t) l_j(t) l_k(t) \quad (1)
$$

where  $g_k(...)$  are the regression coefficients of the model and:

$$
l_i(t) = \sum_{n=1}^{\tilde{N}(t)} \phi_i(n) (RR_{\tilde{N}(t)-n} - RR_{\tilde{N}(t)-n-1}).
$$
 (2)

Straightforwardly, the  $2^{nd}$  NARL is obtained setting the cubic order  $k = 0$  in eq. 1. According to eq. 2, the nonlinear regression in both  $2^{nd}$  NARL and  $3^{rd}$  NARL is performed on the derivative RR series in order to improve the achievement of stationarity within the sliding time window  $W = 90$  sec [4]. The NAR model corresponds to a  $2^{nd}$  NARL model when  $\alpha = 0$ .

For each of the three models, the Newton-Raphson procedure is used to maximize the local log-likelihood defined in [3] in order to estimate the unknown time-varying parameter set  $\xi(t)$ . The optimal order is estimated by means of the Akaike Information Criterion (AIC) and of the point process model goodness-of-fit applied to a subset of the data [3]. Such a goodness-of-fit is based on the Kolmogorov-Smirnov (KS) test. Autocorrelation plots are also considered to test the independence of the model-transformed intervals [3].

The quantitative tools, illustrated in the next sections, are selectively associated to each model. Indeed, the IDLE is exclusively extracted from the  $3^{rd}$  NARL model, whereas the instantaneous bispectral features are exclusively derived from both the  $2^{nd}$  NARL and NAR models. However as preliminary step, when a NARL formulation is considered, the Laguerre deconvolution [7] should always be used to obtain the correspondent long-memory NAR kernel  $\gamma_n(...)$ from the fitted coefficients  $g_n(...)$  (see eq. 1).

# *A. Instantaneous Bispectral Features from* 2 nd *NARL and NAR models*

Three further steps are needed to estimate the instantaneous spectrum and bispectrum and related features of the considered heartbeat series.

- 1) from  $\gamma_n(...)$  find the extended kernels  $\gamma'_n(...)$  [4];
- 2) compute the Fourier transforms  $\Gamma'_n(\ldots)$  of the extended kernels  $\gamma'_n(\ldots);$
- 3) starting from the Fourier transforms of the extended NAR kernels,  $\Gamma'_1(f_1)$  and  $\Gamma'_2(f_1, f_2)$ , compute the Wiener-Volterra Input-Output kernels of order  $p$ ,  $H_p(f_1, \ldots, f_n)$  using the following recursive relationships [10]:

$$
\sum_{p=\text{mid}(q)}^{q} \sum_{\sigma \in \sigma_q} H_p(f_{\sigma(1)}, \dots, f_{\sigma(r)}, \omega_{\sigma(r+1)} + f_{\sigma(r+2)}, \dots, f_{\sigma(q-1)} + f_{\sigma(q)}) \times \Gamma'_1(f_{\sigma(1)}) \cdots \Gamma'_1(f_{\sigma(r)}) \times \Gamma'_2(f_{\sigma(r+1)}, f_{\sigma(r+2)}) \cdots \Gamma'_2(f_{\sigma(q-1)}, f_{\sigma(q)}) = 0
$$
\n(3)

where  $q$  is a given integer representing the kernel order,  $mid(q) = \lceil q/2 \rceil, r = 2p - q$  and  $\sigma_q$  is the permutation set of  $\mathbb{N}_q$ .

Given the  $\Gamma_1'(f_1)$  term, it is possible to compute the timevarying parametric (linear) autospectrum [4] of the original (i.e., non-derivative) heartbeat series:

$$
Q(f,t) = 2(1 - \cos(\omega))S_{xx}(f,t)H_1(f,t)H_1(-f,t)
$$
 (4)

where  $S_{xx}(f, t) = \sigma_{RR}^2$ . By integrating eq. (4) in each frequency band, the power within the very low frequency (VLF = 0.01-0.05 Hz), low frequency (LF =  $0.05$ -0.15 Hz), and high frequency ( $HF = 0.15{\text -}0.5$  Hz) ranges is computed.

Moreover, given the  $\Gamma_2'(f_1, f_2)$  term, it is possible to define the instantaneous bispectrum as reported in detail in [4]. Since the bispectrum presents several symmetry properties that divide the  $(f_1, f_2)$  plane in symmetric zones, for a real signal the bispectrum is uniquely defined by its values in the triangular region of computation  $\Omega$ ,  $0 \le f_1 \le$  $f_2 \leq f_1 + f_2 \leq 1$ . Within this region, it is possible to estimate several features. Specifically, the bispectral invariant [1],  $P(a, t)$  represents the phase of the integrated bispectrum along the radial line with the slope equal to  $a$  (with mean  $P(a, t)$  and variance  $\sigma_{P(a,t)}$ ). The series  $P(a, t)$  results translation, dc-level, amplification, and scale invariant. The parameter  $0 < a \leq 1$  is the slope of the straight line on which the bispectrum is integrated. The variables  $I_r(a, t)$  and  $I_i(a, t)$  refer to the real and imaginary part of the integrated bispectrum, respectively. The Mean magnitude  $M_{\text{mean}}(t)$  and the phase entropy  $P_e(t)$  [1] have  $n = 0, 1, ..., N-1$ ; L as the number of points within the region  $\Omega$ ,  $\Phi$  the phase angle of the bispectrum, and  $1(.)$  the indicator function which gives a value of 1 when the phase angle  $\Phi$  is within the range of bin  $\Psi_n$ . The mean magnitude of the bispectrum can be useful in discriminating between processes with similar power spectra but different third order statistics. However, it is sensitive to amplitude changes. The *normalized bispectral entropy*  $(P_1(t))$  [1] and the *normalized bispectral squared entropy*  $(P_2(t))$  [1] are also considered (between 0 and 1) along with the sum of logarithmic amplitudes of the bispectrum [1]. As the sympatho-vagal linear effects on HRV are mainly characterized by the LF and HF spectral powers [2], the nonlinerar sympatho-vagal interactions can be evaluated by integrating the bispectrum in the bidimensional combination of frequency bands,  $LL(t)$ ,  $LH(t)$ , and  $HH(t)$ .

## *B. IDLE estimation from* 3 rd *NARL model*

Let us consider a generic  $n$ -dimensional linear system in the form  $y_i = Y(t) p_i$ , where  $Y(t)$  is a time-varying fundamental solution matrix with  $Y(0)$  orthogonal, and  $\{p_i\}$ is an orthonormal basis of  $\mathbb{R}^n$ . The key theoretical tools for determining the IDLE and the whole spectrum of LEs is the continuous QR factorization of  $Y(t)$  [5]:

$$
Y(t) = Q(t)R(t)
$$

where  $Q(t)$  is orthogonal and  $R(t)$  is upper triangular with positive diagonal elements  $R_{ii}$ ,  $1 \le i \le n$ . Then, LEs are formulated as:

$$
\lambda_i = \lim_{t \to \infty} \frac{1}{t} \log ||Y(t)p_i||
$$
  
= 
$$
\lim_{t \to \infty} \frac{1}{t} \log ||R(t)p_i|| = \lim_{t \to \infty} \frac{1}{t} \log ||R_{ii}(t)||.
$$

The cubic NAR model can be rewritten in an M-dimensional state space canonical representation:

$$
r_n^{(k)} = \begin{cases} r_{n-1}^{(k+1)} & \text{if } k < M \\ F\left(r_{n-1}^{(M)}, r_{n-1}^{(M-1)}, \dots, r_{n-1}^{(2)}, r_{n-1}^{(1)}\right) & \text{if } k = M \end{cases}
$$

By evaluating the Jacobian  $J(n)$  over the time series, the LE can be determined using the QR decomposition:

$$
J(n)Q_{(n-1)} = Q_{(n)}R_{(n)}
$$

This decomposition is unique except in the case of zero diagonal elements. Then the LEs  $\lambda_i$  are given by

$$
\lambda_i = \frac{1}{\tau H} \sum_{j=0}^{H-1} \ln R_{(j)ii}
$$



Fig. 1. Instantaneous heartbeat statistics computed from a representative CHF patient (N. 5) and healthy subject (CNT16786) using a  $3^{rd}$  NARL model. In the panels a) and c), the estimated  $\mu_{RR}(t)$  are superimposed on the recorded CHF and CNT series, respectively. The panels b) and d) show the IDLE on the recorded CHF and CNT series, respectively. The panel e) shows the  $IDLE_{Md} - IDLE_{MAD}$  bi-dimensional plane of complexity in which the triangles represent the CHF patients and circles represent the healthy subjects (CNT).

where  $H$  is the available number of matrices within the local likelihood window of duration W, and  $\tau$  the sampling time step. The estimation of the LEs is performed at each time t from the corresponding time-varying vector of parameters,  $\xi(t)$  [5]. This provides us with a time-varying vector,  $\lambda_i(t)$ , able to track the Lyapunov spectrum in continuous time. In this work, the first LE,  $\lambda_1(t)$  is considered as the instantaneous dominant Lyapunov exponent (IDLE) and its median,  $IDLE_{Md}$ , and median absolute deviation,  $IDLE_{MAD}$ , are taken as features. These features are unique in literature and achieved only using NARL point process models.

### III. EXPERIMENTAL RESULTS

In order to further validate the proposed algorithms ability in tracking nonlinearities and complexity, an experimental RR dataset from healthy subjects and patients with Congestive Heart Failure (CHF) was taken into account. The heartbeat dataset was retrieved from a public source: Physionet (http://www.physionet.org/). Each RR time series was artifact-free (upon human's visual inspection and artifact rejection) and lasted about 50 min (small segments of the original over 20 h recordings). These recordings have been taken as landmark for studying complex heartbeat interval dynamics. The models goodness-of-fit were evaluated using the KS distance: the smaller the KS distance, the better the model fit. Each of the three model performed a good fit on all heartbeat series. The KS distance, in fact, is  $< 0.073$ in all cases but one (NAR model fitting a CHF series with  $KS = 0.103$ . For all the considered subjects, nearly all

TABLE I

GROUP STATISTICS OF FEATURES FROM HEALTHY AND CHF SUBJECTS.



healthy subject groups.

of the KS plots and more than 97% of the autocorrelation samples were within the 95% confidence bounds. Regarding the model order, AIC analysis indicated  $p = 6 \sim 8$  and  $q = 1 \sim 2$  for the NAR model,  $p = 3 \sim 4$  and  $q = 2 \sim 3$ for the  $2^{nd}$  NARL model,  $p = 3 \sim 4$  and  $q = 2 \sim 3$ and  $q = 1 \sim 2$  for the 3<sup>rd</sup> NARL. In both NARL models,  $\alpha = 0.2$  was chosen.

The instantaneous nonlinearity and complexity indices coming from the NAR,  $2^{nd}$  NARL, and  $3^{rd}$  NARL were evaluated in terms of statistical differences between healthy and CHF subjects. Such a difference was expressed in terms of p-values from a non-parametric rank-sum test under the null hypothesis that the medians of the two sample groups are equal. All features were calculated instantaneously with a 5 ms temporal resolution considering the median values over the estimated time series for further evaluations. Values are expressed as median and its respective absolute deviation (i.e. for a feature X,  $X = \text{Median}(X) \pm \text{MAD}(X)$  where  $MAD(X) = |X - Median(X)|$ ). In Fig. 1, the RR interval series from one representative CHF and healthy subject is shown along with the related  $\mu_{\rm RR}$ , and *IDLE*. The bottom panel shows a bi-dimensional plane which clearly portrays that the combination of  $IDLE_{Md}$  and  $IDLE_{MAD}$  is able to separate the two groups in terms of complexity.

Concerning the standard and bispectral features, the  $2^{nd}$ 

NARL give improved discrimination in 8 of the 15 cases. On average, the CHF patients show significantly lower  $\mu_{\rm RR}$ and  $\sigma_{\rm BR}$ . In addition, they show lower nonlinear interactions between the sympathetic and parasympathetic systems, evaluated by means of  $LL$ ,  $LH$  and  $HH$ . Concerning the complexity features from the  $3^{rd}$  NARL,  $IDLE$  results confirm lower complexity on time series from the pathological population with respect to the healthy subjects, although only  $IDLE_{MAD}$  shows results indicating significant higher complexity of cardiovascular dynamics in healthy patients.

## IV. CONCLUSIONS

We have shown how, within a single point-process paradigm, it is possible to characterize heartbeat nonlinear and complex dynamics by using instantaneous estimations of both higher order spectral features (e.g. the bispectrum), and Lyapunov exponents (e.g. the IDLE index). The use of the discrete Laguerre expansions on quadratic and cubic autoregressive Wiener-Volterra models gives several advantages, such as long-term memory and improved performances. The reported results demonstrate that the proposed point process models are able to track the autonomic-mediated short-term cardiovascular control dynamics in both patients and healthy subjects, further characterizing the inherent nonlinearity and complexity of the system. Importantly, results revealed reduced nonlinear interactions between the sympathetic and parasympathetic systems, as well as lower complex dynamics, in CHF patients. These findings are in agreement with the current literature, whereby cardiovascular disorders affect complexity and variability, and may lead to serious pathological events such as heart failure [11].

#### **REFERENCES**

- [1] K. Chua, V. Chandran, U. Acharya, and C. Lim, "Application of higher order statistics/spectra in biomedical signals–a review," *Medical engineering & physics*, vol. 32, no. 7, pp. 679–689, 2010.
- [2] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C. Lim, and J. Suri, "Heart rate variability: a review," *Medical and Biological Engineering and Computing*, vol. 44, no. 12, pp. 1031–1051, 2006.
- [3] R. Barbieri et al., "A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability," *American Journal of Physiology*, vol. 288, no. 1, p. H424, 2005.
- [4] G. Valenza, L. Citi, E. Scilingo, and R. Barbieri, "Using laguerre expansion within point-process models of heartbeat dynamics: A comparative study," in *Proceeding of IEEE-EMBC*, 2012, pp. 29–32.
- [5] L. Citi, G. Valenza, and R. Barbieri, "Instantaneous estimation of high-order nonlinear heartbeat dynamics by lyapunov exponents," in *Proceeding of the IEEE-EMBC*, 2012, pp. 13–16.
- [6] Z. Chen, E. N. Brown, and R. Barbieri, "Characterizing nonlinear heartbeat dynamics within a point process framework," *Biomed. Eng., IEEE Trans. on*, vol. 57, no. 6, pp. 1335–1347, 2010.
- [7] V. Marmarelis, "Identification of nonlinear biological system using laguerre expansions of kernels," *Ann. Biomed. Eng.*, vol. 21, 1993.
- [8] A. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, p. e215, 2000.
- [9] A. Armoundas *et al.*, "A stochastic nonlinear autoregressive algorithm reflects nonlinear dynamics of heart-rate fluctuations," *Annals of biomedical engineering*, vol. 30, no. 2, pp. 192–201, 2002.
- [10] J. M. Le Caillec and R. Garello, "Nonlinear system identification using autoregressive quadratic models," *Signal processing*, vol. 81, no. 2, pp. 357–379, 2001.
- [11] P. Schwartz and G. De Ferrari, "Sympathetic-parasympathetic interaction in health and disease: abnormalities and relevance in heart failure," *Heart failure reviews*, pp. 1–7, 2010.