Local Holder regularity-based modeling of RR intervals

Olivier Barrière, Jacques Lévy Véhel

Local Hölder regularity-based modeling of RR intervals

Olivier Barrière, Jacques Lévy-Véhel
Projet APIS INRIA Saclay - Île-de-France
Parc Orsay Université, 4 rue J. Monod, 91893 Orsay Cedex France
olivier.barriere@inria.fr, jacques.levy-vehel@inria.fr

Abstract

We analyze the local regularity of RR traces from ECG through the computation of the so-called Hölder exponents. These exponents are at the basis of multifractal analysis, which has been shown to be relevant in the study of RR data. While multifractal analysis yields a global picture of the (statistical) distribution of regularity, we focus here on its time evolution. We show that this evolution is strongly correlated with the signal itself, a feature that seems to have remained unnoticed so far. We use this fact to build realistic synthetic RR traces.

1 Introduction and Motivations

Fractal analysis was developed in view of studying complex irregular objects. Numerous natural phenomena, in particular in physics, biology and medicine, have been shown to exhibit a fractal behaviour. The study of associated models, when available, has lead to significant progress in the understanding and control of these phenomena, for instance in turbulence analysis, non-linear growth, chemical catalysis and wave propagation in irregular media.

Fractal analysis has also been applied with some success in the biomedical field. An example is provided by the study of ECG. ECG and signals derived from them, such as RR intervals, are an important source of information in the detection of various pathologies, including *e.g.* congestive heart failure and sleep apnea. The fractality of these data has been reported in numerous works over the past years. Several fractal parameters, such as the box dimension, have been found to correlate well with the condition of the heart in certain situations ([11, 13]).

More precise information than just a fractional dimension is provided by *multifractal analysis*. Multifractal analysis describes the local regularity at each point, and how it is distributed, either geometrically (leading to the computation of the so-called *Haussdorf multifractal spectrum*), or statistically (in this case, one computes the *large deviation* and *Legendre* multifractal spectra). The statistical approach is by far the most common in applied areas, in particular because the Haussdorf multifractal spectrum is almost impossible to estimate from real data.

Multifractal analysis is useful when the local regularity varies wildly from point to point, and when this irregularity, or its global distribution, holds relevant information on the studied phenomenon. A multifractal behaviour (*i.e.* erratic variations of the local regularity) often emerges as the result of the complex interactions of a large number of elements, each of which acting in a relatively simple way. Such a situation is frequently encountered in the bio-medical field. It is thus no surprise that many studies has been devoted to multifractal analysis of ECG and related signals. In the specific case of RR intervals, several studies has shown a ubiquitiuous multifractal behaviour. In addition, the multifractal spectrum has been found to correlate well with the condition of the heart ([6, 9, 5]). Roughly speaking, one observes two main phenomena:

- On average, the local regularity of healthy RR is smaller than the one in presence of, *e.g.*, congestive heart failure. In other words, pathologies *increases* the local regularity.
- Healthy RR have much more variability in terms of local regularity: Congestive heart failure *reduces* the range of observed regularities.

This results may be traced back to the fact that congestive heart failure is associated with profound abnormalities in both the sympathetic and parasympathetic control mechanisms that regulate beat-to-beat variability ([5]).

These findings are similar to ones made in other areas. Indeed, it is worthwhile noticing the following dichotomy about multifractal phenomena: "natural" multifractality seems to always be a positive quality that constitutes an efficient answer to some functional constraints. Ex-

amples include the organization of the blood and air flows in the lungs, the geometry of tree branches and many more. In contrast, multifractality of artifacts often constitutes an unwanted complication: for instance, it worsens the behaviour of queues in TCP traffic and makes financial assets management more complex.

A precise view on the mechanisms leading to multifractality is important if one wants to understand the purposes it serves and how it will be modified in response to external changes or in case of abnormal behaviour. Multifractal models are largely yet to be developed for RR intervals ([2]). As a preliminary step toward this goal, we study in this work the time-evolution of the local regularity. In other words, instead of estimating, through indirect ways (as is done in multifractal analysis), the distribution of the local regularity, we compute it at each point of the signal. As mentioned above, the local regularity varies rapidly in time. This may seem to preclude its estimation on sampled data. However, recent stochastic models and associated identification procedures now allow to obtain meaningful results even on extremely erratic data. Specifically, we will use in the following the *multifractional Brownian motion* (mBm) as a model for RR traces, and use an estimation method based on *Generalized Quadratic Variations* (GQV).

Computing the time evolution of the local regularity gives far more information than the sole multifractal spectrum. Indeed, the latter may be computed from the former, while the reverse is not true. In addition, inspecting the variation of local regularity yields new insights which cannot be deduced from a multifractal spectrum, since all time-dependent information is lost on a spectrum. In particular, we find that *the evolution of the local regularity is strongly (negatively) correlated with the RR signals*. This fact seems to have remained unobserved so far. It prompts for the development of new models that would account for the fact that, when the RR intervals are *larger*, the RR signal is *more irregular*, and vice-versa. This seems to be a long-term goal. In this work, we content ourselves with using our finding to build realistic synthetic RR traces where the local regularity matches this intriguing correlation. In that view, we develop a new mathematical model that goes beyond the usual mBm.

The remaining of this paper is organized as follows. In section 2, we recall some basic facts about local regularity. In section 3, we present the mBm and its main properties, along with the estimation method we shall use. In section 4, we perform the computation of the local regularity of RR traces (obtained from the PhysioNet database), and show empirically that they are correlated with the signals. In section 5, we explain how to construct an extended mBm allowing a functional relation between the process and its local regularity. We use this construction to generate synthetic RR traces in the last section.

2 Local regularity

There are many ways to measure the local regularity of a signal. A popular one, which has both firm theoretical bases and strong intuitive content is to use the the pointwise Hölder exponent. For a stochastic process whose trajectories are continuous and nowhere differentiable, this exponent is defined as the stochastic process $\{\alpha_X(t)\}_{t\in\mathbf{R}}$ given, for every t, by

$$\alpha_X(t) = \sup\left\{\alpha, \limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{|h|^\alpha} = 0\right\}.$$

Let us explain the geometrical meaning of the statement: "$\alpha_X(t)$ allows to measure the local variations of regularity of $X(t)$." Roughly speaking, saying that a function f has exponent α at t_0 means that, around t_0, the graph of f "looks like" the curve $t \mapsto f(t_0) + c|t-t_0|^\alpha$ in the following sense: for any positive ε, there exists a neighbourhood of t_0 such that the path of f inside this neighbourhood is included in the envelope defined by the two curves $t \mapsto f(t_0) + c|t-t_0|^{\alpha-\varepsilon}$ and $t \mapsto f(t_0) - c|t-t_0|^{\alpha-\varepsilon}$, while this property is no longer true for any negative ε (see figure 1). A "large" α means that f is smooth at t_0, while an irregular behaviour of f at t_0 translates into α close to 0.

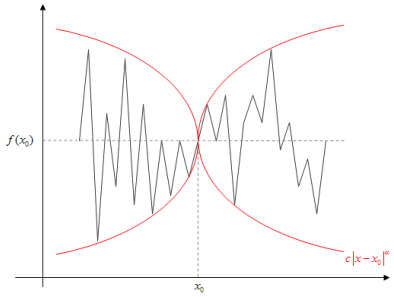

Figure 1. Graphical interpretation of the Hlder regularity of a function f at a point t_0.

3 Multifractional Brownian motion

A popular model for studying rough signals is provided by the so-called fractional Brownian motion (fBm), which was first introduced by Kolmogorov [7] and then studied by Mandelbrot et al. [8].

The one dimensional fBm is the zero-mean Gaussian process $\{B_H(t), t \in \mathbb{R}\}$ with autocorrelation function:

$$R_{B_H}(t,s) = \mathsf{E}\{B_H(t)B_H(s)\} = \frac{\sigma^2}{2}\left[|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right], \quad (1)$$

where $0 < H < 1$ is called the Hurst parameter. (when $H = 1/2$, this is just the well known Brownian motion $B(t)$). It admits the

following integral representation:

$$B_H(t) \propto \int_{-\infty}^{0} \left[(t-s)^{H-1/2} - (-s)^{H-1/2}\right] dB(s) + \int_0^t (t-s)^{H-1/2} dB(s).$$

The first order increments of the sampled fBm (the *fractional Gaussian noise (fGn)* are defined as: $G_H(j) = B_H(j) - B_H(j-1),\ j \in \mathbb{Z}$. $G_H(k)$ is a stationary process. When $H = 1/2$ (Brownian motion), the increments are independent. The case $H \neq 1/2$ is very different, as $G_H(k)$ displays strong correlations in these situations. More precisely, one finds that the autocorrelation function $\rho(k) = \mathsf{E}(G_H(j)G_H(j+k))$ is asymptotically of the order of $|k|^{2H-2}$ for large lags k. In particular, for $1/2 < H < 1$, the correlations are not summable, *i.e.* $\sum_0^{+\infty} |\rho(k)| = +\infty$ This property is usually referred to as *long-range dependence*. Long range dependence is a crucial feature that makes fBm a good model for phenomena ranging from TCP traffic traces to financial records, physiological data, and natural terrain synthesis.

An important property of fBm is related to its local regularity, as measured by the pointwise Hölder exponent α, whose definition is recalled in the previous section. One may prove that, for the fBm, this exponent is with probability one equal to H at all times. Thus, fBms with large (*i.e.* close to 1) H are smoother than those with H close to 0. Finally, recall that the fBm is $H-$self-similar, *i.e.*, for all $a > 0$, $B_H(a\boldsymbol{x}) \stackrel{d}{=} a^H B_H(\boldsymbol{x})$, where $\stackrel{d}{=}$ means the equality of all its finite-dimensional probability distributions.

The stationary-increments property of fBm is useful because it allows to simplify the analysis. However, most real world phenomena do not share this property. In order to obtain realistic models, it is necessary to consider more complex processes, which have non stationary increments of any order. One of the simplest fractal models that falls into this category in the *multifractional Brownian motion* (mBm) [10], which was introduced to overcome certain limitations of the fBm. The major difference between the two processes is that, unlike fBm, the almost sure Hölder exponent of mBm is allowed to depend on time, a useful feature when one needs to model processes whose regularity varies, as is the case for most bio-medical signals. More precisely, a function $H(t)$ replaces the Hurst exponent in the case of mBm. This is important in many situations where one needs a fine modeling. For instance, the long term correlations of the increments of fBm decay as $k^{(2H-2)}$, resulting in long range dependence when $H > 1/2$. In this respect, fBm is "degenerate" since H rules both the high frequencies related to the Hölder regularity and the low frequencies related to the long term dependence. It is thus not possible to have at the same time an irregular local behavior (implying H close to 0) and long range memory (implying $H > 1/2$). fBm is not adapted to model processes which display both those features, such as RR signals or certain highly textured images with strong global organization, as are *e.g.* MR brain images. In contrast, mBm is perfectly adapted in this case. We give now a formal definition of mBm.

Definition 3.1. *The following random function is called multifractional Brownian motion with functional parameter $H(t)$, where $H : [0,\infty) \to [a,b] \subset (0,1)$ is a C^1 function:*

$$W_{H(t)}(t) = \int_{-\infty}^{0} [(t-s)^{H(t)-1/2} - (-s)^{H(t)-1/2}] dW(s) + \int_0^t (t-s)^{H(t)-1/2} dW(s).$$

The mBm is a zero mean Gaussian process, whose increments are in general neither independent nor stationary. One can show that the increments of mBm display long range dependance for all admissible non-constant regularity functions $H(t)$ ([3]).

The main feature of mBm is that its Hölder exponent may be easily prescribed: At each point t_0, it is equal to $H(t_0)$ with probability one. Thus, mBm allows to describe phenomena whose regularity evolves in time/space.

Another important property of mBm is that it is *asymptotically locally self-similar*. Basically, this means that, at each t, there exists an fBm with exponent $H(t)$ which is "tangent" to the mBm. In other words, a path of an mBm can be seen as the lumping of infinitesimal portions of fBm-s with well-chosen exponents

Figure 2 shows two paths of mBm with a linear function $H(t) = 0.2 + 0.6t$ and a periodic $H(t) = 0.5 + 0.3\sin(4\pi t)$. One clearly sees how the regularity evolves in time.

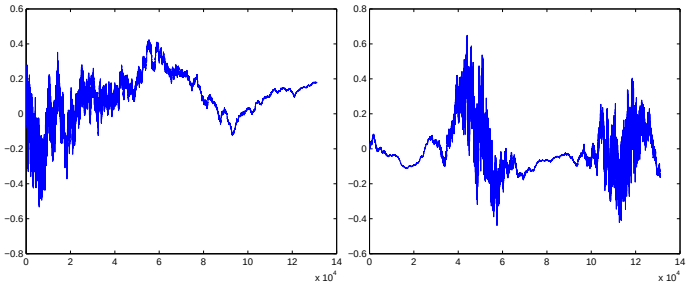

Figure 2. mBm paths with linear and periodic H functions

4 Estimating the local regularity of RR traces

The mBm is, within a multiplicative constant, completely determined by the functional parameter H. Sometimes, it is useful to consider a slightly more general process defined as $G(t)W_{H(t)}(t)$, where G is a deterministic smooth (*e.g.* C^1) function. This allows to take into account trends, periodic behaviours and various other features of natural phenomena. In order to use this modified mBm as a model for time series such as RR intervals, one needs to estimate H and G. This is done in the following way.

The parameter H is estimated through the computation of the so-called "Generalized Quadratic Variations" (GQV). They are defined as follows. For $N \geq 1$, let $\left\{X\left(\frac{p}{N}\right);\ p \in \{0,\ldots,N-1\}\right\}$ denote a discretized trajectory of the process $\{X(t)\}_{t\in[0,1]}$. The GQV are the quantities:

$$\tilde{V}_N(t) = \sum_{p \in \tilde{\nu}_N(t)} \left(X\left(\frac{p}{N^\delta}\right) - 2X\left(\frac{p+1}{N^\delta}\right) + X\left(\frac{p+2}{N^\delta}\right)\right)^2$$

where $\tilde{\nu}_N(t) = \left\{p \in \mathbb{N}; 0 \leq p \leq N-2 \text{ and } \left|t - \frac{p}{N^\delta}\right| \leq N^{-\gamma}\right\}$.

It has been proven in [1] that for any mBm $W_{H(t)}$, under some technical assumptions on δ and γ, one has almost surely:

$$\lim_{n\to\infty} \frac{1}{2\delta}\left((1-\gamma) - \frac{\log \tilde{V}_N(t)}{\log N}\right) = H(t)$$

.

This leads to a simple estimator

$$\hat{H}_N(t) = \frac{1}{2\delta}\left((1-\gamma) - \frac{\log \tilde{V}_N(t)}{\log N}\right)$$

This is very efficient when ones observes a standard mBm, *ie* when $G \equiv 1$, but presents in general a discrepancy $\Delta_G(t)$:

$$\Delta_G(t) = -\frac{\log G(t)}{\delta \log N}$$

In order to get rid of this bias, a classical approach consists in using a least squares linear regression of $\log \tilde{V}_N$ *versus* $\log N$. From the slope α_1 and the intercept α_0, one deduces a second estimator of H and an estimator of G

$$\begin{aligned} \hat{H}_{reg}(t) &= -\frac{\alpha_1(t) - (1-\gamma)}{2\delta} \\ \hat{G}_{reg}(t) &= \exp(-\alpha_0(t)/2); \end{aligned}$$

This regression effectively eliminates the bias, but at the price of a noticeable augmentation of the variance. We thus used a hybrid method allowing to decrease the variance without increasing the bias. This technique consists in aligning the temporal mean of the first estimator on the one of the second and is described in [4]. The bias is then eliminated and the variance remains small.

$$\hat{H}_{N\_reg} = \hat{H}_N - < \hat{H}_N > + < \hat{H}_{reg} >$$

where $< \hat{H} >= \frac{1}{N}\sum_{i=0}^{N-1} \hat{H}(t_i)$

Finally, we use the discrepancy between $\hat{H}_{N\_reg}$ and $\hat{H}_N$ to estimate G

$$\hat{G}_{N\_reg}(t) = N^{\delta\left(\hat{H}_{N\_reg} - \hat{H}_N\right)}$$

Figure 3 illustrates typical estimations of H on mBm paths (the theoretical regularity is in green and the estimated one is in blue).

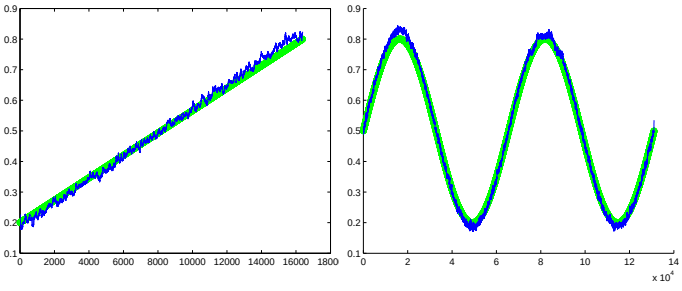

Figure 3. Estimation of the local regularity of mBm paths. Left: linear H function. Right: periodic H function

5 Regularity of RR traces

We obtained real 24-hour interbeat (RR) interval time series from the PhysioNet database [12]. These were derived from long-term ECG recordings of adults between the ages of 20 and 50 who have no known cardiac abnormalities and typically begin and end in the early morning (within an hour or two of the subject's awakening. For each signal, composed of around $100,000$ points, we estimated its Hölder exponent (in green) using the method based on GQV described in section 4.

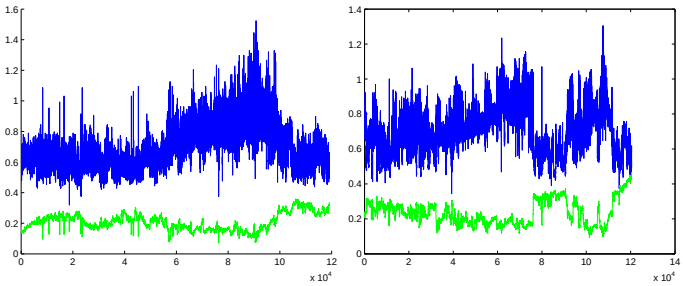

Figure 4. Estimation of the local regularity of RR interval time series

As one can see it on figure 4, there is a clear negative correlation between the value of the RR interval and its local regularity: when the blue curves move up, the green ones tend to move down. In other words, slower heartbeats (higher RR values) are typically more irregular (smaller Hölder exponents) than faster ones.

6 Improving the mBm-based model by allowing the local regularity to depend on the amplitude

In order to account for the striking feature that, at time instants where the RR intervals are larger, the heartbeat is typically also more irregular, we need to refine the modeling based on mBm. Indeed, while the mBm allows to tune the regularity at each time, it does so in an "exogenous" manner. This means that the value of H and of W_H are independent. In light of the previous paragraph, a better model for RR time series requires to define a modified mBm where the regularity would be a function of W_H at each time. We shall call this process a Self-Regulating Multifractional Process (SRMP). It is defined as follows. We give ourselves a deterministic smooth one-to-one function g ranging in $(0,1)$, and we seek a process X such that, at each t, $\alpha_X(t) = g(X(t))$ almost surely. It is not possible to write an explicit expression for such a process. Rather, we rely on a fixed point approach that we now describe briefly. See [4] for more details. We start from an mBm W_H as given in definition **3.1**, with an arbitrary function H (for instance a constant). At the second step, we set $H = g(W_H)$. We then iterate this process, *i.e.* we compute a new W_H with this updated H function, and so on. One may prove that these iterations will almost surely converge to a well-defined SRMP X with the desired property, namely the regularity of X at any given time t is equal to $g(X(t))$.

For such a process, there is a functional relation between the amplitude and the regularity. However, this does not allow a precise control of the Hölder exponent. Let us explain this on an example. Take for definiteness $g(t) = t$ for all t. Then, a given realization might result in a low value of X at, say, $t = 0.5$ and thus high irregularity at this point, while another realization might give a large $X(0.5)$ resulting in a path that is smooth at 0.5. See figure 5 for an example of this fact.

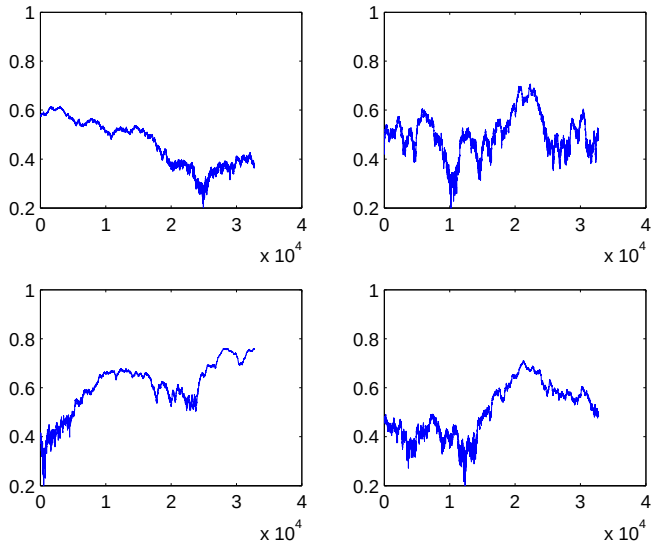

Figure 5. Paths of SRMPs with $g(Z) = Z$

In order to gain more control, we generalize the definition of an SRMP as follows. We first define a "shape function" s, which is a deterministic smooth function. Then, at each step, we compute W_H, and set $H = g(W_H + ms)$, where m is a positive number. The function s thus serves two purposes. First, it allows to tune the shape of X: when m is large, X and s will essentially have the same form. Second, because of the first property, it allows to decide where the process will be irregular and where it will be smooth. Figure 6 displays an example of SRMP with controlled shapes.

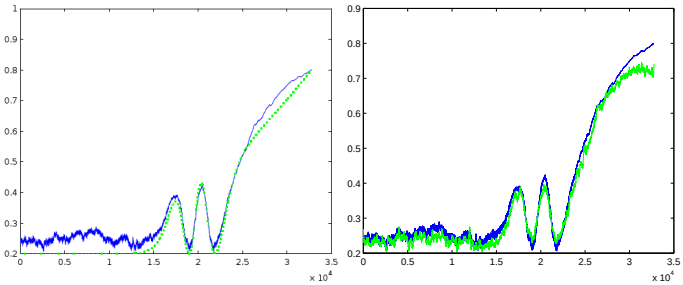

Figure 6. Left: SRMP with $g(Z) = Z$ (blue), and controlling shape function (green). Right: same SRMP (blue) and estimated Hlder exponent (green).

7 RR traces synthesis

Our model for RR traces relies on the following ingredients:

- An "s" function, that describes the overall shape of the trace, and in particular the nycthemeral cycle.
- A g function whose role is to ensure that the correct relation between the heart rate and its regularity is maintained at all times.

We estimate s from our data in the following way : For each RRi time series, we plotted histograms of both the signal and its exponent, and modeled it as a sum of two Gaussians, as represented on figure 7

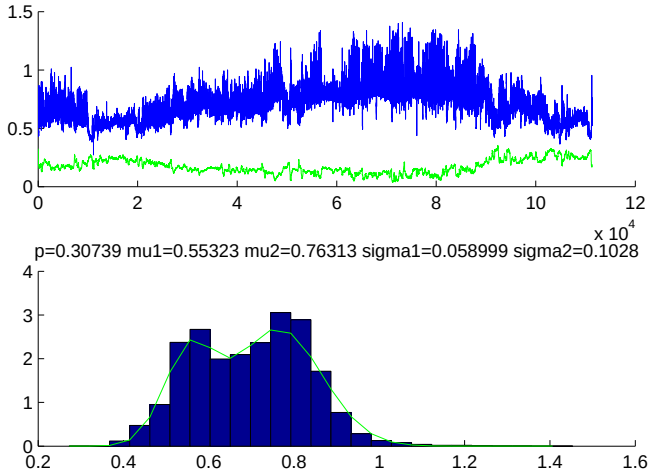

Figure 7. Histogram of RRi time series, modeled as a sum of two Gaussians

We then inferred from theses signals the following shape function, represented on figure 8, based on splines and parameterized by :

- D_n, duration of the night: $D_n \in [6, 10]$
- D_m, duration of the beginning of the measure : $D_m \in [2, 4]$
- D_s, duration of the sleeping phase : $D_s \in [0.5, 1.5]$
- D_a, duration of the awakening phase : $D_r \in [0.5, 1.5]$
- RRi_d, mean interbeat interval during the day : $RRi_d \in [0.6018, 0.7944]$
- RRi_n, mean interbeat interval during the night : $RRi_n \in [0.7739, 1.0531]$

randomly chosen in their respective intervals, with uniform probability.

The g function is estimated in the phase space. More precisely, we plot, for each trace, the value of H as a function of the RR interval. Representing all these graphs on a single plot, we get a histogram, as in figure 9.

We then extract the ridge lines of this histogram, seen as a surface in the (RR, α) plane (see figure 9). It is seen that this ridge line is roughly a straight line, that we fit using least squares minimization in order to obtain an equation of the form $\alpha = g(RR) = aRR + b$.

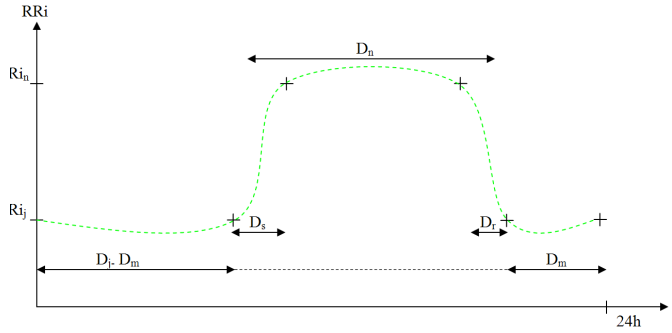

Figure 8. Shape function of RR intervals

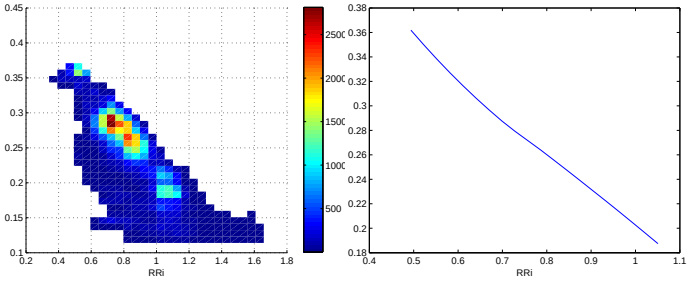

Figure 9. Histogram in the phase space

The last step is to synthesize an SRMP with shape function s and regularity function g, as explained in the previous section. Paths obtained in this way are shown on figure 10. Compare with the graphs shown on figure 4 displaying true RR traces.

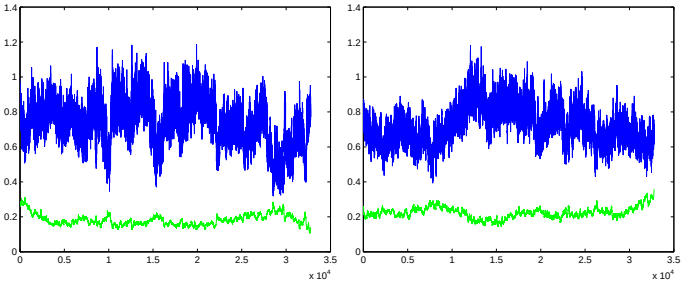

Figure 10. Two forged RR intervals based on SRMP (blue) and estimated regularity (green).

8 Conclusion and future work

A fine analysis of the local regularity of RR traces reveals the intriguing property that the Hölder exponent and amplitude are typically related in a linear way at each time instant. We have build a phenomenological model to account for this fact, introducing a new stochastic process termed Self-Regulating Multifractional Process.

Further work will concentrate on trying to explain, in physiological terms, the origin of the relation between α and the RR value. We will also try to detect whether this relation is affected by various pathologies. Finally, the impact of this relation on the multifractal spectrum of RR intervals will be investigated.

References

[1] AYACHE, A. AND LÉVY VÉHEL, J. (2004). On the identification of the pointwise Hölder exponent of a generalized multifractional Brownian motion. *SPA* **111,** 119–156.

[2] L. A. N. Amaral, A. L. Goldberger, P. C. Ivanov, H. E. Stanley *Modeling heart rate variability by stochastic feedback*, Computer Phys. Comm. **121:122**, 1999, 126–128

[3] AYACHE, A. AND COHEN, S. AND LÉVY VÉHEL, J. (2000). The covariance structure of multifractional Brownian motion, with application to long range dependence. *ICASSP*.

[4] BARRIÈRE, O. (2007). Synthese et estimation de mouvements Browniens multifractionnaires et autres processus a regularite prescrite. Definition du processus auto-regule multifractionnaire et applications. *Ph.D. Thesis, Univ. Nantes*.

[5] A. L. Goldberger, L. A. N. Amaral, J. M. Hausdorff, P. C. Ivanov, C.K. Peng, H. E. Stanley, *Fractal dynamics in physiology: Alterations with disease and aging*, PNAS **99**, 2002, 2466–2472.

[6] C. Ivanov, L.Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, Z. R. Struzik, H. E.Stanley, *Multifractality in human heartbeat dynamics*, Nature **399**, June 1999.

[7] KOLMOGOROV A.N. (1940). Wienersche Spiralen und einige andere interessante Kurven in Hilbertchen Raume, *Doklady*, **26**, 115–118.

[8] MANDELBROT, B.B. AND VAN NESS, J. (1968). Fractional Brownian motion, fractional noises and applications. *SIAM Rev.* **10,** 422–437.

[9] M. Meyers, O. Stiedl and B. Kerman, *Discrimination by multifractal spectrum estimation of human heartbeat interval dynamics*, Fractals **11-2**, June 2003.

[10] PELTIER, R.F. AND LÉVY VÉHEL, J. (1995). Multifractional Brownian motion: definition and preliminary results. http://www-rocq1.inria.fr/fractales/index.php?page=publications

[11] PETERS, R. (1999). The Fractal Dimension of Atrial Fibrillation:A New Method to Predict Left Atrial Dimension from the Surface Electrocardiogram, Cardiology, **92**(1), 17–20.

[12] www.physionet.org

[13] TU, C., ZENG, Y. AND YANG, X. (2004) Nonlinear processing and analysis of ECG data, Technology and Health Care **12**(1),1–9.