



OPEN

Analysis of temporal correlation in heart rate variability through maximum entropy principle in a minimal pairwise glassy model

Elena Agliari¹, Francesco Alemanno^{2,3}, Adriano Barra^{2,4}, Orazio Antonio Barra^{5,6}✉, Alberto Fachechi^{2,4}, Lorenzo Franceschi Vento⁶ & Luciano Moretti^{7,8}

In this work we apply statistical mechanics tools to infer cardiac pathologies over a sample of M patients whose heart rate variability has been recorded via 24 h Holter device and that are divided in different classes according to their clinical status (providing a repository of labelled data). Considering the set of inter-beat interval sequences $\{r(i)\} = \{r_1(i), r_2(i), \dots\}$, with $i = 1, \dots, M$, we estimate their probability distribution $P(r)$ exploiting the maximum entropy principle. By setting constraints on the first and on the second moment we obtain an effective pairwise (r_n, r_m) model, whose parameters are shown to depend on the clinical status of the patient. In order to check this framework, we generate synthetic data from our model and we show that their distribution is in excellent agreement with the one obtained from experimental data. Further, our model can be related to a one-dimensional spin-glass with quenched long-range couplings decaying with the spin-spin distance as a power-law. This allows us to speculate that the $1/f$ noise typical of heart-rate variability may stem from the interplay between the parasympathetic and orthosympathetic systems.

Heart-rate variability (HRV) analysis constitutes a major tool for investigating the mechanisms underlying the complex and chaotic cardiac dynamics as well as for identifying general features discriminating the clinical status of patients^{1,11,14,15,16,18,19}. To this aim a fundamental observable is the RR series $\mathbf{r} = \{r_1, r_2, \dots\}$, where r_n is the temporal distance between the n -th and the $(n + 1)$ -th R peaks in a ECG recording (see Fig. 1 left panel). Several approaches have been carried out in the past in order to address the HRV analysis from this observable (see e.g.^{17,37,38}). For instance, in^{1,2} the problem of classification of heart failures via time-series analysis is translated into a search for clusterization in a high-dimensional space.

Interestingly, the intrinsic variability in heart rate ultimately stems from the interplay of the sympathetic and the parasympathetic nervous system. In this work, exploiting statistical inference approaches, we aim to unveil any signature of this underlying autonomic neural regulation. To this scope we will study HRV in the temporal and in the frequency domain, and at different levels of aggregation (in the higher one the sample is made of all available data, in the lower one we build different sub-samples pertaining to patients displaying a different clinical status: healthy, suffering from cardiac decompensation, suffering from atrial fibrillation).

The statistical inference approach we adapt to the present case of study is the *maximum entropy framework*³ recently developed and applied to pitch sequences for capturing melodic styles⁴. In general, the maximum

¹Dipartimento di Matematica "Guido Castelnuovo", Sapienza Università di Roma, P. le A. Moro, 00185 Roma, Italy. ²Dipartimento di Matematica e Fisica "Ennio De Giorgi", Università del Salento, Via per Arnesano, 73100 Lecce, Italy. ³Institute of Nanotechnology, National Research Council (CNR-NANOTEC), Campus Ecotekne, Via Monteroni, 73100 Lecce, Italy. ⁴Istituto Nazionale di Fisica Nucleare (INFN), Campus Ecotekne, Via Monteroni, 73100 Lecce, Italy. ⁵Department of Environmental Engineering, University of Calabria (INICAL-DIAM), 87035 Cosenza, Italy. ⁶Politecnico Internazionale "Scientia et Ars" (POLISA), Largo Intendenza, 89900 Vibo, Valentia, Italy. ⁷Department of Cardiology "C. & G. Mazzoni", Hospital (APH), Via degli Iris, 63100 Ascoli-Piceno, Italy. ⁸Luciano Moretti is deceased. ✉email: orazioantonio.barra@unical.it

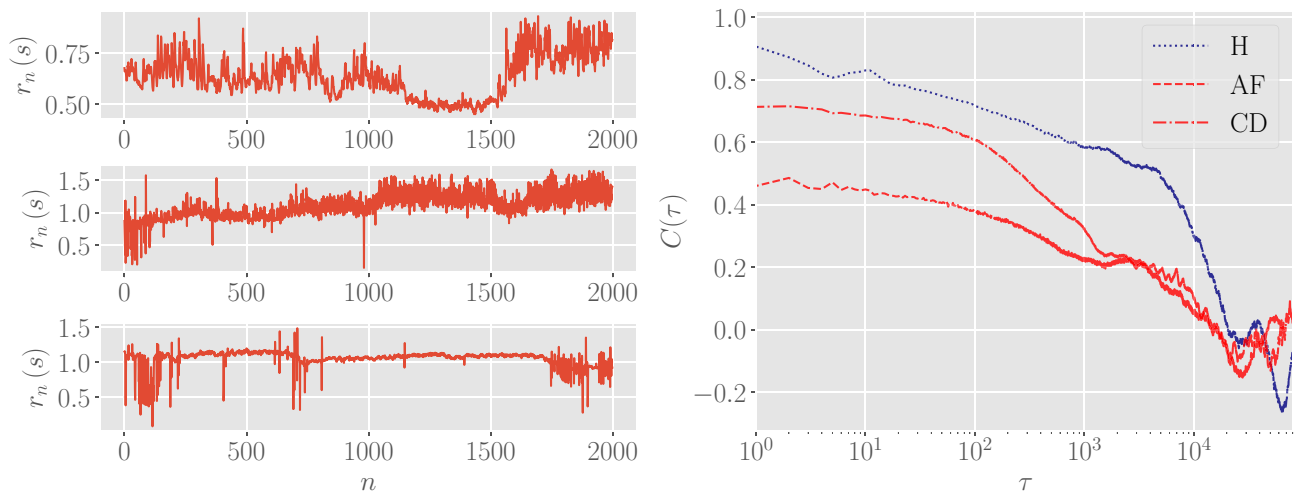


Figure 1. Left: examples of the bare RR time series for a single patient for each class; the window depicted is restricted to the first 2000 beats. Right: examples of autocorrelation functions for a single patient for each class. The dotted blue line refers to a healthy patients, while red are patients with AF (dashed curve) and CD (dash-dotted line). Notice that, in any case, the autocorrelation remains positive over several orders of magnitude and this is a consequence of the scale-free behavior of the series power spectral density (see also Fig. 7, upper panel).

entropy principle has been widely applied to biological problems, ranging from the cellular scales of neuronal¹⁶ and immunological settings⁵, up to social scales of bird's flocks⁸ and ant's communications⁹. By this technique, we search for the minimal-structure probabilistic model compatible to our data; more precisely, we consider the family of probability distributions $P(\mathbf{r})$ over the sequences of inter-beat intervals \mathbf{r} whose lowest momenta match the empirical ones and, among all the elements of this family we select the one corresponding to the maximum entropy. As a direct consequence of the definition of entropy in terms of the logarithm of the probability distribution $P(\mathbf{r})$ over the inter-beat sequences \mathbf{r} , this approach returns an exponential family $P(\mathbf{r}) \sim e^{-H(\mathbf{r})}$, where $H(\mathbf{r})$ can be interpreted as a cost function (or Hamiltonian in a physical jargon). By requiring a match on the first two moments only (i.e., by requiring that the theoretical average and two-point correlation provided by the model are quantitatively consistent with the empirical ones), $H(\mathbf{r})$ results in a pairwise (r_n, r_m) cost-function, as standard in Physics.

Of course, recovering the complex structure hidden in RR series by a relatively simple pairwise model has several advantages: on the one hand, the low number of parameters prevents from over-fitting, on the other hand, the inferred cost-function can be framed in a statistical mechanics context (see e.g.,⁵⁻⁹) and we can therefore rely on several powerful techniques and on a robust Literature. In particular, we will show that, despite its simplicity, such a pair-wise model is able to capture the complex nature of the temporal correlation between beats which emerges experimentally; in fact, the coupling between two beat-intervals r_n and $r_{n+\tau}$ turns out to be long-range (i.e., displays a power-law decay with the distance τ) and frustrated (i.e., the couplings between two beats can be positive and negative). In a statistical-mechanical jargon, this system is referred to as a two-body spin-glass with power-law quenched interactions.

Remarkably, frustration in couplings, which is a key feature of spin-glasses, means the existence of competitive driving forces and it is natural to look at this emerging feature in our model as the hallmark of the interplay between the parasympathetic and orthosympathetic systems (indeed, while the first one tends to increase the distance between RR peaks, i.e., to lower the heart rate¹⁸, the latter tends to decrease it²⁰). We speculate that these competing interactions may be responsible for the well-known $1/f$ noise shown by HRV^{12,13,21}: spin-glasses typically display a chaotic dynamics²²⁻²⁴ spread over several timescales²⁵ and their power spectrum density is power-law³²⁻³⁴. In fact, here we show that the autocorrelation in the $\{r_n\}$ series decays in the beat number as n^{-1} and its related power spectrum decays in frequency as f^{-1} . Incidentally, we notice that variables whose fluctuations display $1/f$ noise are widespread, ranging from inorganic (e.g., condensed²⁶, granular²⁷, etc.) to organic (e.g. in DNA sequences²⁸, membrane channels²⁹) matter and, even broadly, in Nature (e.g. ranging from earthquakes³⁵ to off-equilibrium flows of current through resistors³⁶, to the whole self-organized criticality^{30,31}).

Results

Summary of experimental data. In this Section, we give some details about the data and the quantities considered in our analysis. This research has been accomplished as a part of the Project *MATCH* (Mathematical Advanced Tools to Catch Heart-Rate-Variability), a scientific cooperation among Ascoli Piceno Hospital (APH), POLISA, University of Calabria and University of Salento: Holter recordings have been acquired at APH and their data-base built at POLISA during the period 2016–2019.

The database is made of ECG recordings on $M = 348$ patients, wearing an Holter device for nominal 24 h. From these recordings we extract the RR series

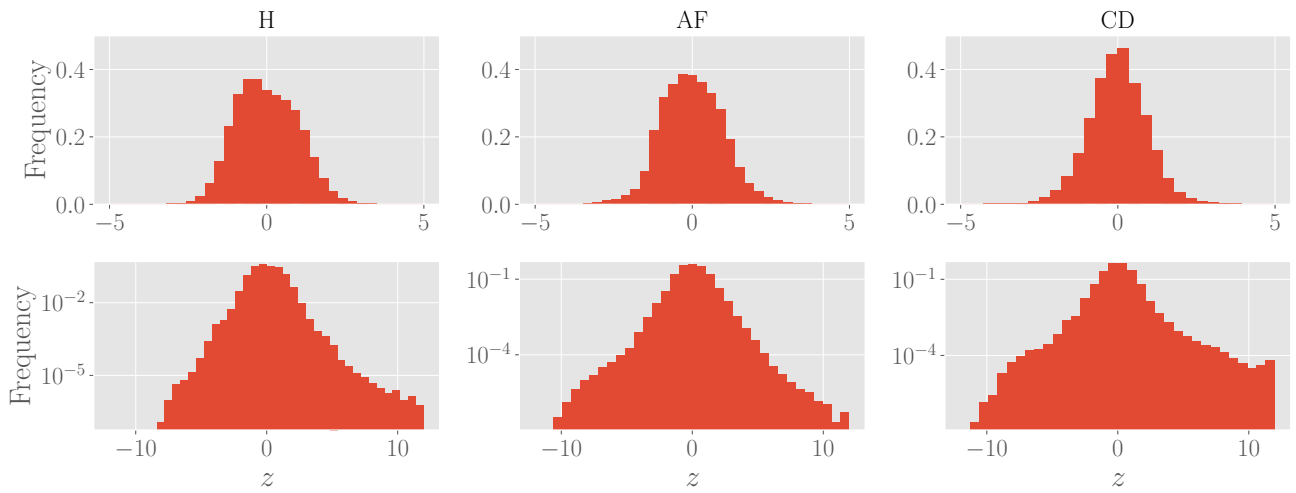


Figure 2. Histograms of the standardized values $\{z(i)\}$ divided by classes: left panels are built by collecting data from healthy patients, middle panels are built by collecting data from patients suffering from atrial fibrillation and right panels are built by collecting data from patients suffering from cardiac decompensation. In the first row, we reported relative frequencies in the natural scale, while the second row we reported relative frequencies in the logarithmic scale.

$$\{\mathbf{r}(i)\}_{i=1,\dots,M} = \{r_n(i)\}_{\substack{i=1,\dots,M \\ n=1,\dots,N_i}} \tag{1}$$

where i labels the patient and n labels the number of beats in each sequence (which is order of 10^5 and depends on the patient). Patients belong to three classes, according to their clinical status: healthy individuals (H), individuals with atrial fibrillation (AF) and individuals with congestive heart failure (hereafter simplified as *cardiac decompensation*) (CD). Their number is $M_H = 149$, $M_{AF} = 139$, and $M_{CD} = 60$, respectively; of course, $M = M_H + M_{AF} + M_{CD}$. In Fig. 1 (left) we show examples of the series $\mathbf{r}(i)$ for three patients belonging to the different classes.

In order to make a meaningful comparison of the variability among the RR series $\mathbf{r}(i)$ of different patients, we standardize them with respect to their temporal mean and standard deviation, so that the study of HRV is recast in the study of fluctuations of the standardized RR series around the null-value. More precisely, we introduce

$$z_n(i) = \frac{r_n(i) - \langle \mathbf{r}(i) \rangle}{\text{std}[\mathbf{r}(i)]}, \text{ for } n = 1, \dots, N \tag{2}$$

or, in vectorial notation,

$$\mathbf{z}(i) = \frac{\mathbf{r}(i) - \langle \mathbf{r}(i) \rangle}{\text{std}[\mathbf{r}(i)]}, \tag{3}$$

where we defined

$$\langle \mathbf{r}(i) \rangle = \frac{1}{N_i} \sum_{n=1}^{N_i} r_n(i), \quad \langle \mathbf{r}^2(i) \rangle = \frac{1}{N_i} \sum_{n=1}^{N_i} r_n^2(i), \quad \text{std}[\mathbf{r}(i)] = \sqrt{\langle \mathbf{r}^2(i) \rangle - \langle \mathbf{r}(i) \rangle^2}. \tag{4}$$

The raw histograms for the standardized inter-beat intervals in the three classes of patients are shown in Fig. 2: notice that the frequency distributions exhibit heavy-tails.

We consider the points in the standardized RR series as random variables sampled by a hidden stochastic process, in such a way that the value of $z_n(i)$ at a given step n depends in principle on all the values $\{z_m(i)\}_{m < n}$ taken in the previous steps $m < n$ since the beginning of sampling. From this perspective, a meaningful observable to look at is the auto-correlation function at a distance τ , defined as

$$C(i, \tau) = \frac{1}{N_i - \tau} \sum_{n=1}^{N_i - \tau} (z_n(i) - \langle z(i) \rangle_+) (z_{n+\tau}(i) - \langle z(i) \rangle_-), \tag{5}$$

where $\langle z(i) \rangle_+ = \frac{1}{N_i - \tau} \sum_{n=1}^{N_i - \tau} z_n(i)$ and $\langle z(i) \rangle_- = \frac{1}{N_i - \tau} \sum_{n=\tau}^{N_i} z_n(i)$. Given the standardization over the whole segment $[1, N_i]$, as long as $\tau \ll N_i$, we expect that $\langle z(i) \rangle_+$ and $\langle z(i) \rangle_-$ are both close to zero and shall be neglected in the following (indeed, we checked that this is the case, since $\langle z(i) \rangle_+, \langle z(i) \rangle_- \sim 10^{-15} \div 10^{-17}$). Then, the auto-correlation function we measure simply reduces to

$$C(i, \tau) = \frac{1}{N_i - \tau} \sum_{n=1}^{N_i - \tau} z_n(i) z_{n+\tau}(i). \quad (6)$$

Some examples of the autocorrelation function for patients of the three classes are reported in the right plot of Fig. 1, where we stress that the autocorrelation is non-null over a large τ window and its shape is patient-dependent. The last property stems from the $1/f$ noise shown by HRV which, in turns, yields to dominant contributions in the low-frequency domain of the RR series.

Finally, we introduce a further average operation, this time on the sample of patients, namely, we define

$$\mathbb{E}_{class}(\mathbf{z}) = \frac{1}{M_{class}} \sum_{i \in class} \mathbf{z}(i), \quad (7)$$

$$\mathbb{E}_{class}(\mathbf{z}^2) = \frac{1}{M_{class}} \sum_{i \in class} \mathbf{z}^2(i), \quad (8)$$

$$\mathbb{E}_{class}(C(\tau)) = \frac{1}{M_{class}} \sum_{i \in class} C(i, \tau). \quad (9)$$

where $class \in \{H, AF, CD\}$ and with “ $i \in class$ ” we mean all the indices corresponding to patients belonging to a certain class. In the following we will consider the vectors \mathbf{z} as random variables sampled from an unknown probability distribution $P_{class}^{true}(\mathbf{z})$, which we will estimate by the probability distribution $P_{class}(\mathbf{z})$ characterized by a minimal structure and such that its first and second moments are quantitatively comparable with $\mathbb{E}_{class}(\mathbf{z})$, $\mathbb{E}_{class}(\mathbf{z}^2)$ and $\mathbb{E}_{class}(C(\tau))$, respectively.

Discussion on the model and on the inferential procedure. Our atomic variable is the sequence $\{z_1, z_2, \dots, z_N\}$ and, as anticipated above, we denote with $P(\mathbf{z})$ the related probability distribution emerging from the inferential operations on the sample of experimental data. The Shannon entropy $\tilde{H}[P(\mathbf{z})]$ associated to $P(\mathbf{z})$ is

$$\tilde{H}[P(\mathbf{z})] = - \int d\mathbf{z} P(\mathbf{z}) \ln P(\mathbf{z}). \quad (10)$$

According to the maximum entropy principle, we look for the distribution $P(\mathbf{z})$ that maximizes $\tilde{H}[P(\mathbf{z})]$ and such that its moments match those evaluated experimentally, in particular, here we choose to apply the constraints on the one-point and two-points correlation function that is, $\mathbb{E}_{class}(\mathbf{z})$, $\mathbb{E}_{class}(\mathbf{z}^2)$ and $\mathbb{E}_{class}(C(\tau))$, respectively. To lighten the notation hereafter these moments shall be referred to simply as, respectively, $\mu^{(1)}$, $\mu^{(2)}$ and $C(\tau)$, without specifying the class. In fact, the inferential procedure works analogously regardless of the class, the latter affecting only the quantitative value of the parameters occurring in $P(\mathbf{z})$. Constraints are set via Lagrange multipliers $(\lambda_0, \lambda_1, \lambda_2, \lambda_\tau)$ in such a way that the problem is recast in the maximization of the functional

$$\begin{aligned} \tilde{H}_{\lambda_0, \lambda_1, \lambda_2, \lambda_\tau}[P(\mathbf{z})] &= \tilde{H}[P(\mathbf{z})] + \lambda_0 \left(\int d\mathbf{z} P(\mathbf{z}) - 1 \right) \\ &+ \lambda_1 \left(\sum_{n=1}^N \int d\mathbf{z} P(\mathbf{z}) z_n - N \mu^{(1)} \right) + \lambda_2 \left(\sum_{n=1}^N \int d\mathbf{z} P(\mathbf{z}) z_n^2 - N \mu^{(2)} \right) \\ &+ \sum_{\tau=1}^N \lambda_\tau \left(\sum_{n=1}^{N-\tau} \int d\mathbf{z} P(\mathbf{z}) z_n z_{n+\tau} - (N - \tau) C(\tau) \right), \end{aligned} \quad (11)$$

where integration is made over \mathbb{R}^N . Note that, while the derivation with respect to λ_1 , λ_2 and λ_τ ensure, respectively, the agreement between the theory and the experiments at the two lowest orders, i.e. the temporal average $\mu^{(1)}$, the second moment $\mu^{(2)}$ and the auto-correlation function $C(\tau)$, λ_0 guarantees that $P(\mathbf{z})$ is normalized, so that $P(\mathbf{z})$ is a probability distribution function. In the asymptotic limit of long sampling ($N \rightarrow \infty$) and under a stationarity hypothesis (see⁴ for a similar treatment and the section dedicated to Methods for the proof), the solution of the extremization procedure, returning the probability of observing a certain sequence \mathbf{z} , is given by

$$P(\{z_n\}_{n=1}^\infty) = \frac{1}{Z} \left(\prod_{n=1}^\infty P_0(z_n) \right) \exp \left(\sum_{n=1}^\infty \sum_{\tau=1}^\infty J(\tau) z_n z_{n+\tau} + h \sum_{n=1}^\infty z_n \right), \quad (12)$$

where h and $J(\tau)$ can be estimated from available data (*vide infra*). Here, P_0 is the $\mathcal{N}(0, 1)$ distribution and plays the role of prior for the variable z_n , the parameter $J(\tau)$ represents the pairwise interaction between elements at non-zero distance τ in the series (notice that each element occurs to be coupled to any other), and the parameter h represents the bias possibly affecting the single value in the sequence (and it is expected to be zero as we standardized the RR series). The factor Z plays here as a normalization constant, like the partition function in the statistical mechanics setting¹⁰. Notice that the interaction between two elements r_n and r_m ($n > m$) depends on the distance $\tau = n - m$, but not on the particular couple considered. This stems from a “stationary hypothesis”, meaning that one-point and two-point correlation functions calculated on a segment spanning $O(\tau \ll N)$

elements along the series are approximately the same and since the starting time of sampling is arbitrary, we get that $J(n, m) = J(m - n)$.

The standard inference setup for the model parameters is based on a *Maximum (log-)Likelihood Estimation* (MLE), i.e. the maximization of the function

$$\mathcal{L}(\mathbf{J}, h | \mathcal{D}) = -\frac{1}{M} \sum_{\mathbf{z} \in \mathcal{D}} \log P(\mathbf{z} | \mathbf{J}, h), \quad (13)$$

where \mathcal{D} is the time-series database (of a given class) and where we made clear the dependence of P on the model parameters. However, such an approach requires the computation of the whole partition function Z , which is numerically hard in this case. Then, we chose to adopt as objective function for the inference procedure the *pseudo-(log-)likelihood* function⁴:

$$\mathcal{L}(\mathbf{J}, h | \mathcal{D}) = -\frac{1}{M} \sum_{\mathbf{z} \in \mathcal{D}} \log P(z_{L+1} | \{z_n\}_{n=1}^L), \quad (14)$$

that is, given L observations in a fixed time-series \mathbf{z} , we maximize the conditional probability to observe the value z_{L+1} at the successive time step (i.e., we are introducing a cut-off in the interaction range). Further, we make two main modifications with respect to the standard pseudo-likelihood approach: (i) in order to use the entire available time-series in our database, we also adopt a window average procedure; (ii) we add regularization terms in order to prevent divergence for the model parameters. A detailed discussion is reported in Appendix 4.2. Our objective function is therefore given by

$$\mathcal{L}^{(\text{reg})}(\mathbf{J}, h | \mathcal{D}) = \frac{1}{M} \sum_{\mathbf{z} \in \mathcal{D}} \left[-\frac{1}{2(N-T)} \sum_{n=T}^{N-1} (z_{n+1} - h - \sum_{\tau=1}^T J(\tau) z_{n+1-\tau})^2 - \frac{\lambda}{2} h^2 - \frac{\lambda}{2} \sum_{\tau=1}^T f(\tau) J(\tau)^2 \right], \quad (15)$$

where T is the largest τ we want to consider (namely T must be larger than the maximal decorrelation time), λ is the regularization weight and $f(\tau)$ is a temporal regularizer that prevents the elements of $J(\tau)$ to get too large for large τ (see Appendix 4.2 for a detailed description).

The inference method allows us to determine the values of the parameters $J(\tau)$ and h as well as their uncertainties $\sigma_{J(\tau)}$ and σ_h . As for the parameter h , due to series standardization, its value, evaluated over the different classes, is expected to be vanishing (this is indeed the case as it turns out to be $h \sim 10^{-3}$ with a related uncertainty of the same order). As for the pairwise couplings, we find that for all the classes considered, $J(\tau)$ is significantly non-zero only for relatively small values of τ , with a cutoff at $T \sim 10^2$, and, for a given $\tau < T$, the coupling does not display a definite sign, that is, for pairs $(z_n, z_{n+\tau})$ and $(z_m, z_{m+\tau})$ at the same distance τ the related couplings can be of opposite signs. These results are shown in Fig. 3: in the left column we reported the inferred $J(\tau)$ with the associated uncertainties for all τ , and in each panel in the right we reported the frequency distributions for the first values of τ as examples.

In order to study how decorrelation of RR intervals takes place, it is interesting to study how the interaction vector $J(\tau)$ (regardless of its sign) vanishes as the delay time τ increases. We found that the long delay time behavior of the magnitude (i.e. disregarding the signs and oscillatory characteristics) of the interactions is well-described by a power law of the form

$$|J(\tau)|_{\text{leading}} \sim A \cdot \tau^{-\beta}. \quad (16)$$

We thus fitted the tails of the inferred $J(\tau)$ with this trial function (the range of fitted points is chosen in order to maximize the adjusted R^2 score). In Table 1, we report best-fit parameters, the adjusted R^2 score and the reduced χ^2 . It is interesting to note that the scaling parameter β is around 1, meaning that, for each of the three classes, the leading behavior of the interactions at large τ is $\sim 1/\tau$ (as we will see later, the same scaling also characterizes the power spectral density in the Fourier domain, see Fig. 7). In the upper row of Fig. 4, we depicted with red circles the results of the inference procedure (once taken their absolute values), while the best fit of the general trend is represented with dashed black lines. In the lower row of the same figure, we also reported the residuals of the experimental values with respect to the best fit (normalized to the corresponding uncertainty). Apart from the first few points in the AF and CD cases, we see that the residuals are distributed in a range of at most $2\sigma_{J(\tau)}$ (where $\sigma_{J(\tau)}$ is the standard deviation at each τ point), and, in particular, for sufficiently long τ (where oscillations are softened), experimental values are always contained in the range $[-\sigma_{J(\tau)}, \sigma_{J(\tau)}]$, implying that the leading behavior of the delayed interaction is well-captured by $\sim 1/\tau$ noise both in our datasets as well as in the model's prediction.

To summarize, here we highlight the main results.

- Couplings can be both positive and negative (see Fig. 3), defining the heart as a complex glassy system.
- The coupling magnitude decays in τ as a power-law whose leading order is $\sim 1/\tau$ (see Fig. 4).
- The coupling magnitude displays a sharp scaling $1/\tau$ solely in healthy patients, while for the remaining patients it displays a *bump* in the short time-scales (see Fig. 3 and the residual plots in Fig. 4).
- By fitting data via Eq. (16), we obtain estimates for the power-law exponent β (as reported in Table 1): interestingly, different classes are associated to different best-fit values of β , in such a way that classification of cardiac failures via HRV via this route seems possible.

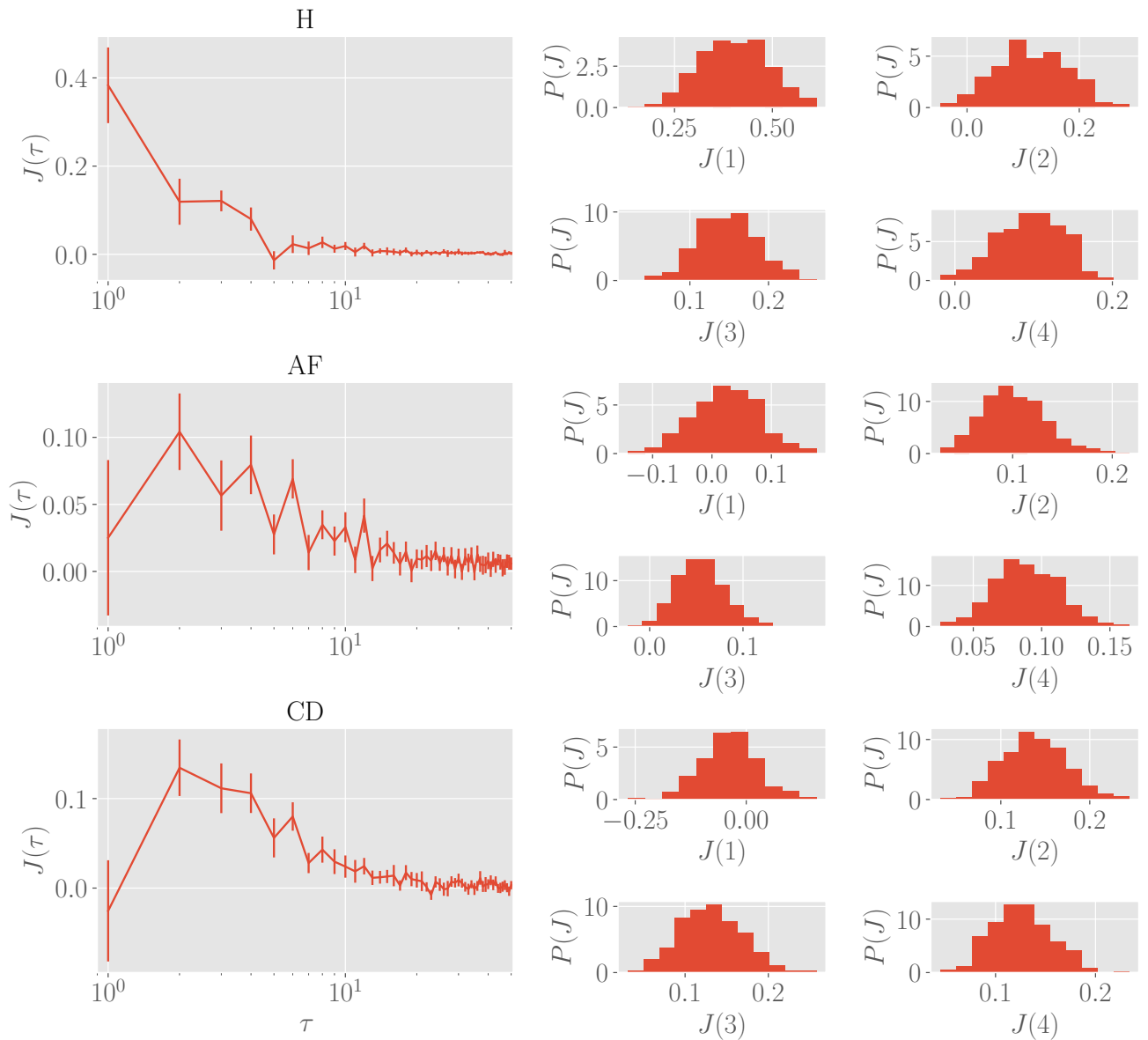


Figure 3. Inference results for delayed interactions. Left column: the plots show the results of the inference procedure (distinguishing between the clinical status) for the first 50 τ s. Right: frequency distribution of the J s for some selected values of τ (i.e. $\tau = 1, 2, 3, 4$). In both cases, the statistics consists in $M = 500$ different realizations of the $J(\tau)$ which are realized by randomly extracting different mini-batches, each with size $n = 20$. We stress that some frequency distributions present tails on negative values of J for some τ . This means that frustrated interactions are also allowed, implying that the system is fundamentally complex, i.e. a *glassy hearth*.

Class	A	β	\bar{R}^2	χ^2/DOF
H	0.38 ± 0.04	1.41 ± 0.04	0.75	0.33
AF	0.20 ± 0.03	0.96 ± 0.05	0.79	0.42
CD	0.37 ± 0.05	1.2 ± 0.05	0.78	0.46

Table 1. Best fit values regarding the scaling reported in Eq. (16). For each class, we report the best-fit parameters A and β , as well as the adjusted R^2 and the reduced χ^2 scores quantifying the fit goodness.

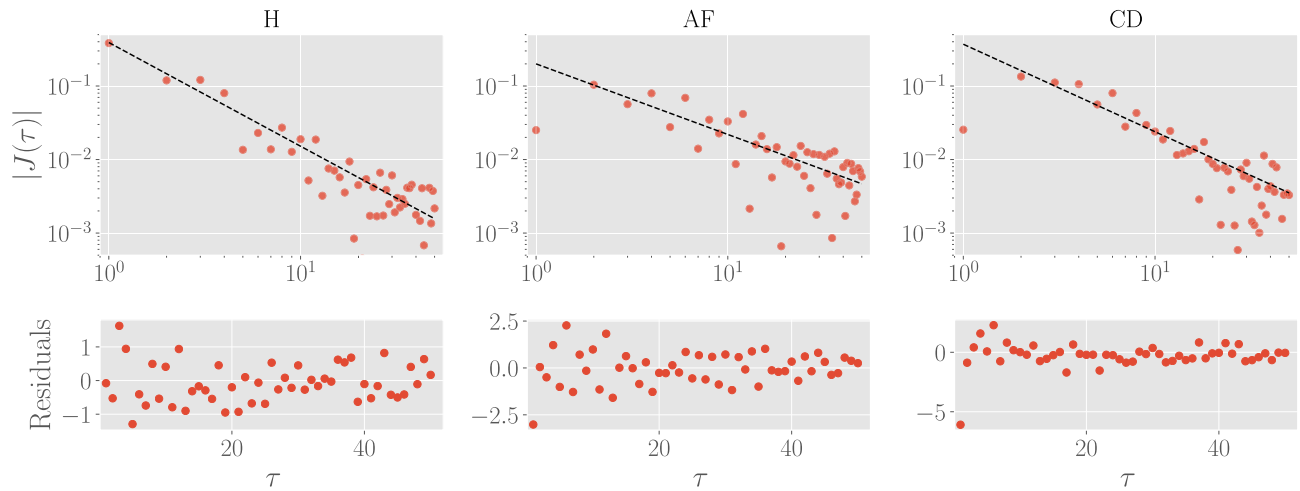


Figure 4. Leading behavior of magnitude of the delayed interactions. In the upper panel, we reported the absolute value of the delayed interactions $J(\tau)$ and the relative best fit. In the lower panel, we reported the residuals (normalized by the uncertainty at each point τ) of the experimental data with respect to the best fit function. We stress that, even if the interactions $J(\tau)$ are far from the fitting curve (in the log-log scale, see first row), they are compatible within the associated uncertainties, as remarked by the residual plots.

Discussion on the model and on the generalization procedure. Once the model and the related parameters are inferred for each of the three classes, we can use the original sequences $\{\mathbf{z}\}$ to generate synthetic sequences $\{\tilde{\mathbf{z}}\}$ of length $N - T$. The procedure followed to get the synthetic sequence is briefly described hereafter, while an extensive explanation is provided in App. 4.3.

For any class, we consider our estimate for $J(\tau)$, along with the estimate $\sigma_{J(\tau)}$ of its uncertainty, and we build the noisy estimate for $J(\tau)$, that is $\tilde{J}(\tau) = J(\tau) + \delta J(\tau)$ where $\delta J(\tau) = \eta \sigma_{J(\tau)}$ and η is a $\mathcal{N}(0, 1)$ random variable. Next, taken a certain $\{\mathbf{z}\}$, we convolve it with $\tilde{J}(\tau)$ and this returns $\{\tilde{\mathbf{z}}\}$. Of course, due to the initial standardization of the RR series, the inference procedure returned a vanishing bias parameter h , hence the synthetic series will also be centered at zero. However, a synthetic sequence is no longer standardized and this is done by hand.

Then, it is natural to compare the synthetic sequences and the experimental ones. We generate a sample of data with the same size of the experimental data available, and we compute the empirical cumulative distribution function for both the experimental and synthetic data in order to compare them: results are reported in Fig. 5. In the first row, we directly compare the experimental (red solid line) and the synthetic (black dashed line) cumulative distributions highlighting an excellent agreement for all the classes. This is then corroborated by checking the probability plots in the same figure (second row): here, the red solid line shows the synthetic cumulative distribution versus experimental cumulative distribution, while the black dashed curve is the identity line. The green regions in the plot are confidence intervals with $p = 0.95$.

Next, we test whether the model is able to effectively capture correlation in the RR series, in particular by comparing experimental auto-correlation functions and their predicted counterparts. However, since autocorrelation functions are individual-dependent, starting from a single (randomly chosen) RR series we generate 100 synthetic series with different realizations of the $\tilde{J}(\tau)$ according to the above mentioned procedure (see also App. 4.3). In this way, we can use our estimation of the uncertainties on $J(\tau)$ in order to give a confidence interval for our predictions. In Fig. 6 we compare the auto-correlation functions for the experimental series and for the synthetic series; for the latter we also highlight the confidence interval with $p = 0.68$. More precisely, we depict the experimental (red solid line) and theoretical (black dashed line) auto-correlation functions and see that the former always fall inside the confidence interval of the re-sampled series (the green region). Thus, we can conclude that our inferred minimal pairwise model is able to effectively capture the temporal autocorrelation in the RR series.

As a final comment, we also looked at the power spectrum density (PSD) of the provided datasets $\{\mathbf{z}\}$ that, as expected (see e.g.,^{12,13,21}), displays the long tail $1/f$ (see Fig. (7), upper panel) and we made the following comparison: for all the patients, we evaluated its PSD and in the region $[10^{-4}, 10^{-2}]$ Heartbeat⁻¹ we fit with a power-law

$$\text{PSD}(f) = \alpha \cdot f^{-\gamma}$$

where $\gamma \sim 1$ and its value is taken as the x -coordinate of that patient in the lower panels of Fig. (7). The corresponding y -value is obtained by calculating the PSD over 100 synthetic RR-series generated by convolution with the empirical series playing as seed and using as value of J the one pertaining to the class the patient belongs to (H, AF, CD); results are in good agreement on the diagonal.

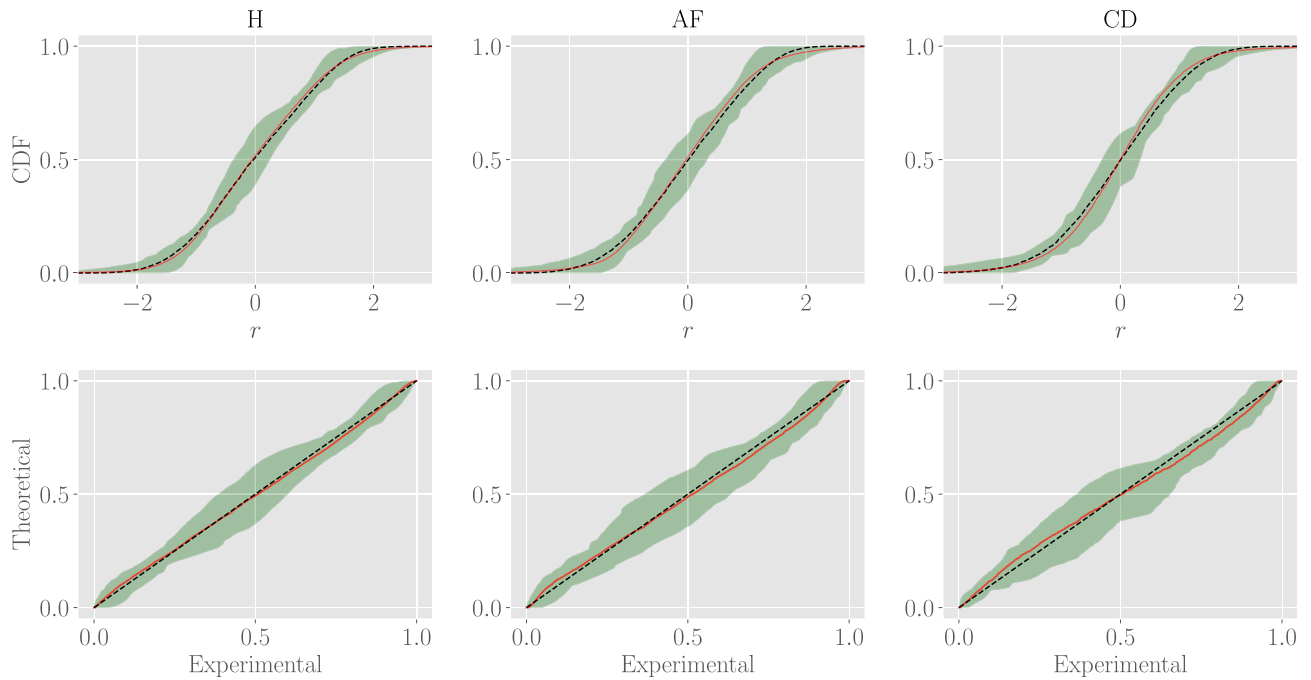


Figure 5. Comparison between posterior distributions for experimental and synthetic data. First row: comparisons between the empirical cumulative distributions for both experimental (solid red lines) and resampled (black dashed lines) populations for all of the three classes. Second row: probability plots for the two populations of data (i.e. empirical *versus* theoretical ones, red solid lines) for all of the three classes. The black solid curves are the identity lines for reference. The green region is the confidence interval with $p = 0.95$.

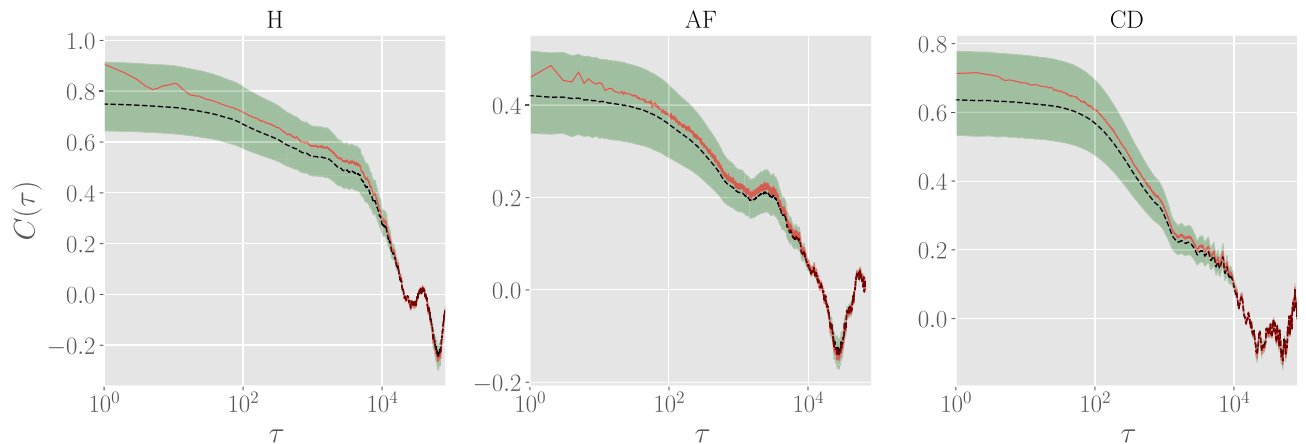


Figure 6. Comparison between autocorrelation functions for experimental and synthetic data. The autocorrelation function for one patient randomly extracted from the experimental data-set (red solid lines) is compared with the median autocorrelation function obtained from the synthetic dataset (black dashed lines). Notice that the former falls in the confidence interval with $p = 0.68$ (green region) of the latter.

Conclusion

Several past studies have highlighted that heart-rate fluctuations, in healthy individuals, exhibit the characteristic $1/f$ noise (see¹² and references therein). Deviations from this behavior can in fact be associated to cardiac pathologies such as atrial fibrillation or congestive heart failure²¹. In this work we tried to deepen the mechanisms possibly underlying this peculiar behavior, both in healthy and in compromised subjects.

To this aim we exploited inferential tools derived from statistical mechanics (i.e. the maximum entropy principle) to construct a probability distribution $P(\mathbf{r})$ characterizing the occurrence of a RR series \mathbf{r} . By requiring that $P(\mathbf{r})$ is minimally structured (i.e., prescribing the maximum entropy) and that $P(\mathbf{r})$ correctly matches the empirical first and second moments, we end up with a probabilistic model analogous to a spin-glass where quenched couplings $J(\tau)$ among spins exhibit frustration and a power-law decay with the distance τ between spin pairs. This kind of system is known to display chaotic dynamics spread over several timescales and the $1/f$ noise. We thus speculate that the presence of competitive driving force are key features for the emergence of the

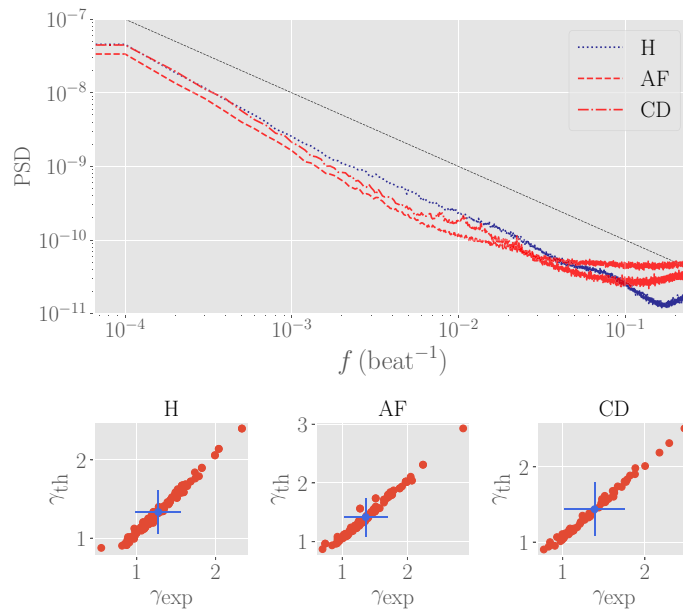


Figure 7. Top: empirical power spectral density (PSD). The dotted blue line refers to a healthy patients, while the red ones refer to patients with AF (dashed curve) and CD (dash-dotted line). The PSD is computed according to the Welch procedure with 50% windows overlap. The black continuous curve is the expected $1/f$ -noise distribution for visual comparison. Bottom: scatter plot for the scaling exponent of the PSD (in the region 10^{-4} e 10^{-2} Hz); in particular, we take the simple average over the synthetic realizations, the red spots are the exponent for the single patient (notice that the uncertainties over the synthetic realization are much smaller and are not visible in the plot), and the blue spot marks the average over all patients (both experimental and synthetic) with the relative uncertainties.

rich phenomenology displayed by heart-rate and we are naturally tempted to identify the two opposite driving forces with the sympathetic and parasympathetic systems.

We stress that our model for $P(\mathbf{r})$ is robustly checked and that tests performed over different classes of patients (displaying a different clinical status) allows us to candidate the exponent β controlling the coupling decay $J(\tau) \sim \tau^{-\beta}$ as an indicator for classify the patient clinical status.

Methods

This Section is devoted to the description of the statistical model and the optimization rules used in our inference procedure.

The two-body statistical mechanical model from maximum entropy principle. As stated in Sec. 2, the probabilistic model we use to frame the analysis of heart-rate variability contained in the standardized RR series $\{z_n\}_{n=1}^N$ emerges as the solution of extremization procedure of the constrained Shannon entropy functional

$$\begin{aligned} \tilde{H}_{\lambda_0, \lambda_1, \lambda_2, J}[P(\mathbf{z})] = & \tilde{H}[P(\mathbf{z})] + \lambda_0 \left(\int d\mathbf{z} P(\mathbf{z}) - 1 \right) \\ & + \lambda_1 \left(\sum_{n=1}^N \int d\mathbf{z} P(\mathbf{z}) z_n - N\mu^{(1)} \right) + \lambda_2 \left(\sum_{n=1}^N \int d\mathbf{z} P(\mathbf{z}) z_n^2 - N\mu^{(2)} \right) \\ & + \sum_{\tau=1}^N \lambda_{\tau} \left(\sum_{n=1}^{N-\tau} \int d\mathbf{z} P(\mathbf{z}) z_n z_{n+\tau} - (N - \tau)C(\tau) \right). \end{aligned} \tag{17}$$

Here, we recall that the first term is the standard Shannon entropy for probability distribution for continuous variables, i.e.

$$\tilde{H}[P(\mathbf{z})] = - \int d\mathbf{z} P(\mathbf{z}) \log P(\mathbf{z}), \tag{18}$$

while the other terms are constraints with Lagrangian multipliers $\lambda_0, \lambda_1, \lambda_2$ and λ_{τ} . The extremization with respect to these parameters leads to the following conditions:

$$\int dz P(\mathbf{z}) = 1, \quad \frac{1}{N} \sum_{n=1}^N \int dz P(\mathbf{z}) z_n = \mu^{(1)}, \quad \frac{1}{N} \sum_{n=1}^N \int dz P(\mathbf{z}) z_n^2 = \mu^{(2)}, \quad \frac{1}{N-\tau} \sum_{n=1}^{N-\tau} \int dz P(\mathbf{z}) z_n z_{n+\tau} = C(\tau), \tag{19}$$

i.e. that the function $P(\mathbf{z})$ is a probability distribution and that the moments up to the second order are captured by the experimental temporal average $\mu^{(1)}$, the temporal standard deviation $\mu^{(2)}$ and the auto-correlation function $C(\tau)$. The extremization with respect to the function P leads to the explicit form of the solution, i.e.

$$\log P(\mathbf{z}) = \lambda_0 - 1 + \lambda_1 \sum_{n=1}^N z_n + \lambda_2 \sum_{n=1}^N z_n^2 + \sum_{\tau=1}^N \lambda_\tau \sum_{n=1}^{N-\tau} z_n z_{n+\tau}, \tag{20}$$

which can be rewritten as

$$P(\mathbf{z}) = \text{cost} \cdot \exp \left(\lambda_1 \sum_{n=1}^N z_n + \lambda_2 \sum_{n=1}^N z_n^2 + \sum_{\tau=1}^N \sum_{n=1}^{N-\tau} \lambda_\tau z_n z_{n+\tau} \right). \tag{21}$$

The constant in the latter equation is computed by using the normalization property of the probability distribution $P(\mathbf{z})$, and it is given by

$$\text{cost}^{-1} = Z = \int dz \exp \left(\lambda_1 \sum_{n=1}^N z_n + \lambda_2 \sum_{n=1}^N z_n^2 + \sum_{\tau=1}^N \sum_{n=1}^{N-\tau} \lambda_\tau z_n z_{n+\tau} \right), \tag{22}$$

where we used the letter Z to make contact with the notion of partition function from the Thermodynamics dictionary. Since the model is essentially Gaussian, we can directly compute the partition function (at least, in formal way) as

$$Z = \int dz \exp \left(\lambda_1 \mathbf{E}^T \cdot \mathbf{z} + \mathbf{z}^T (\lambda_2 \mathbb{I} + \boldsymbol{\lambda}) \mathbf{z} \right) = (-\pi)^{N/2} \det^{-1/2} (\lambda_2 \mathbb{I} + \boldsymbol{\lambda}) \exp(-\lambda_1^2 \mathbf{E}^T (\lambda_2 \mathbb{I} + \boldsymbol{\lambda})^{-1} \mathbf{E}). \tag{23}$$

where $\mathbf{E} = (1, 1, \dots, 1)$ is a N -dimensional vector of ones and we defined the interaction matrix $(\boldsymbol{\lambda})_{n,m} = \sum_{\tau=1}^N \delta_{n,m-\tau} \lambda_\tau$ (which turns out to be an upper triangular Toeplitz matrix with zeros on the main diagonal). Because of this, the determinant of the kernel $\lambda_2 \mathbb{I} + \boldsymbol{\lambda}$ is trivially $\det(\lambda_2 \mathbb{I} + \boldsymbol{\lambda}) = \lambda_2^N$. We can now determine the relation between the temporal average and standard deviation in terms of the model parameters. These relations read as

$$\mu^{(1)} = \left\langle \frac{1}{N} \sum_{n=1}^N z_n \right\rangle_{\text{th}} = \frac{1}{N} \frac{\partial \log Z}{\partial \lambda_1} = -\frac{2}{N} \lambda_1 \mathbf{E}^T (\lambda_2 \mathbb{I} + \boldsymbol{\lambda})^{-1} \mathbf{E}, \tag{24}$$

$$\mu^{(2)} = \left\langle \frac{1}{N} \sum_{n=1}^N z_n^2 \right\rangle_{\text{th}} = \frac{1}{N} \frac{\partial \log Z}{\partial \lambda_2} = -\frac{1}{2\lambda_2} - \frac{\lambda_1^2}{N} \frac{\partial}{\partial \lambda_2} \mathbf{E}^T (\lambda_2 \mathbb{I} + \boldsymbol{\lambda})^{-1} \mathbf{E}, \tag{25}$$

where $\langle \cdot \rangle_{\text{th}}$ is the theoretical average as defined in Eq. 19.

Since \mathbf{z} is temporally standardized, we directly get $\lambda_1 = 0$ and $\lambda_2 = -1/2$. However, we left the former as a free parameter to be inferred and check *a posteriori* that it is consistent with zero. In order to get contact with Physics' dictionary, we rename the Lagrangian multipliers $\lambda_1 = h$ and $\lambda_\tau = J(\tau)$, playing the role of external magnetic field and two-body interactions respectively. Then, the solution of the maximum entropy problem (after some rearrangements of the sum indices) is given by

$$P(\mathbf{z}) = \frac{1}{Z} \left(\prod_{n=1}^N P_0(z_n) \right) \exp \left(\sum_{n=1}^{N-1} \sum_{\tau=1}^{N-n} J(\tau) z_n z_{n+\tau} + h \sum_{n=1}^N z_n \right), \tag{26}$$

where $P_0(z)$ is the Gaussian distribution $\mathcal{N}(0, 1)$ and Z is the partition function.

The inference setup. The determination of the model parameters $J(\tau)$ and h is based on a maximum likelihood approach. As stated in the model description, the usage of the full probability is computationally untractable (because of the high dimensionality of the integral in the partition function), thus we use a maximum pseudo-likelihood approach⁴ (namely a tractable asymptotic correct estimator of the likelihood), in which the fundamental object to be maximized is the conditional probability that, given the observations $\{z_n\}_{n=1}^L$, the successive observation is equal to experimental data z_{L+1} :

$$P(z = z_{L+1} | \{z_n\}_{n=1}^L) = \frac{P(\{z_n\}_{n=1}^{L+1})}{P(\{z_n\}_{n=1}^L)} = \frac{Z^{(L)}}{Z^{(L+1)}} \frac{\left(\prod_{n=1}^{L+1} P_0(z_n) \right) \exp \left(\sum_{n=1}^L \sum_{\tau=1}^{L+1-n} J(\tau) z_n z_{n+\tau} + h \sum_{n=1}^{L+1} z_n \right)}{\left(\prod_{n=1}^L P_0(z_n) \right) \exp \left(\sum_{n=1}^{L-1} \sum_{\tau=1}^{L-n} J(\tau) z_n z_{n+\tau} + h \sum_{n=1}^L z_n \right)}. \tag{27}$$

The second factor is easy to handle with, while the partition function can be evaluated by using the fact that $\int dz P(z | \{z_n\}_{n=1}^L) = 1$, so that we finally have

$$P(z = z_{L+1} | \{z_n\}_{n=1}^L) = \frac{\exp\left(\log P_0(z_{L+1}) + \sum_{\tau=1}^L J(\tau)z_{L+1-\tau}z_{L+1} + h z_{L+1}\right)}{\int dz \exp\left(\log P_0(z) + \sum_{\tau=1}^L J(\tau)z_{L+1-\tau}z + h z\right)}. \tag{28}$$

Notice that eqs. 27–28 are, in fact, approximations given that interactions between past (before $L + 1$) and future (after $L + 1$) events are neglected. Now, since the prior is Gaussian, we can directly integrate the denominator for carrying out a closed form for the conditional probability. Thus, we get

$$P(z = z_{L+1} | \{z_n\}_{n=1}^L) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(z_{L+1} - \sum_{\tau=1}^L J(\tau)z_{L+1-\tau} - h\right)^2\right). \tag{29}$$

The MLE is based on the maximization of this conditional probability, or equivalently of the pseudo log-likelihood, which is composed by quantities of the form

$$\log P(z = z_{L+1} | \{z_n\}_{n=1}^L) = -\frac{1}{2}\left(z_{L+1} - \sum_{\tau=1}^L J(\tau)z_{L+1-\tau} - h\right)^2, \tag{30}$$

where we discarded unessential constant terms. Since we would like to infer the first values of the delayed interaction vector $J(\tau)$ and since the RR time-series have a size which is of the order of 10^5 , it is better to use a sliding-window average approach, whose functioning is ensured by the stationary hypothesis. In this way, we can also perform a temporal average over all a single RR time-series. Supposing we want to infer the first T elements of the delayed interaction $J(\tau)$ (i.e. we truncate long-term correlations) and given the time-series $\{z_n\}_{n=1}^N$ of length N , we define the individual log-likelihood as

$$\begin{aligned} \mathcal{L}(\mathbf{J}, h | \{z_n\}_{n=1}^N) &= -\frac{1}{N-T} \sum_{L=T}^{N-1} \log P(z_{L+1} | \{z_n\}_{n=L-T+1}^L) \\ &= -\frac{1}{2(N-T)} \sum_{L=T}^{N-1} \left(z_{L+1} - h - \sum_{\tau=1}^T J(\tau)z_{L+1-\tau}\right)^2. \end{aligned} \tag{31}$$

In order to prevent the parameters to acquire large values, we also introduce some regularization. For the bias, we simply add a quadratic penalization term: $\mathcal{R}(h) = -\lambda h^2/2$. Concerning the interaction vector, in order to discourage the algorithm to generate spurious correlation for high τ , we introduce a penalization which depends on the delay time τ , i.e.

$$\mathcal{R}(\mathbf{J}) = -\frac{\lambda}{2} \sum_{\tau=1}^T f(\tau)J(\tau)^2. \tag{32}$$

However, in order to ensure not to destroy correlation for interesting values of τ , we adopt a mild regularizer. In all of our tests, we found that a good choice is $f(\tau) = \log^2(1 + \tau)$. Putting all pieces together, we have the regularized individual pseudo log-likelihood

$$\begin{aligned} \mathcal{L}^{(\text{reg})}(\mathbf{J}, h | \{z_n\}_{n=1}^N) &= -\frac{1}{2(N-T)} \sum_{L=T}^{N-1} \left(z_{L+1} - h - \sum_{\tau=1}^T J(\tau)z_{L+1-\tau}\right)^2 \\ &\quad - \frac{\lambda}{2} h^2 - \frac{\lambda}{2} \sum_{\tau=1}^T f(\tau)J(\tau)^2. \end{aligned} \tag{33}$$

The whole pseudo-likelihood is the average over the set \mathcal{D}_c time-series in each given class (where $c \in \{H, AF, CD\}$):

$$\mathcal{L}^{(\text{reg})}(\mathbf{J}, h | \mathcal{D}_c) = \frac{1}{M_c} \sum_{z \in \mathcal{D}_c} \mathcal{L}^{(\text{reg})}(\{z_n\}_{n=1}^N | \mathbf{J}, h), \tag{34}$$

where M_c is the number of examples in the class. By adopting a standard gradient descent (GD) approach, we can derive the following optimization rules:

$$\delta J(\tau) = \sum_{L=T}^{N-1} \Delta_L z_{L+1-\tau} - \lambda f(\tau)J(\tau), \tag{35}$$

$$\delta h = \sum_{L=T}^{N-1} \Delta_L - \lambda h, \tag{36}$$

where

$$\Delta_L = \frac{1}{N-T} (z_{L+1} - h - \sum_{\tau=1}^T J(\tau) z_{L+1-\tau}). \quad (37)$$

In order to speed up the inference procedure, we use a AdaGrad³⁹ adaptation method for the gradient descent rules (3536). Since we want a uncertainties estimation for the coupling matrix $J(\tau)$, we proceed in the following way: better than to realize a single delayed interaction for the whole database (for each class), we minimize the pseudo log-likelihood to M random subsets of cardinality n of the database of each class (i.e. the gradients are averaged with respect to this minibatch) and then let the inferential algorithm converge towards a fixed point. Then, we compute the mean values and the standard deviation with respect to this M realizations of the interaction vector $J(\tau)$. In our analysis, we choose $M = 500$ different inferential procedures for subset of cardinality $n = 20$.

Finally, we notice that the inferential approach followed here can also be seen as a (least squares) regression, where the regressors are all the past inter-peak durations, i.e., $z_L = \sum_{\tau} J(\tau) z_L - \tau + \text{Gaussian noise}$. Along this line, other examples of statistical mechanics formulation include, for instance, a gaussian version of the Kinetic Ising Model (see e.g.,^{40–44}).

Generation of synthetic data. Having an estimate for the interaction vector $J(\tau)$ and for the bias h , we can use them to generate new synthetic sequences. From Eq. (29), we see that the conditional probability $P(z = z_{L+1} | \{z_n\}_{n=1}^L)$ is maximized if we take

$$z_{L+1} = \sum_{\tau=1}^L J(\tau) z_{L+1-\tau} + h. \quad (38)$$

Therefore, once the first T components of $J(\tau)$ and the bias parameter h are inferred, we can generate synthetic series $\{\tilde{z}_n\}$ as follows. First, we take the first T points in a given experimental time-series, then we convolve it with $J(\tau)$ and add the bias, i.e.

$$\tilde{z}_1 = \sum_{\tau=1}^T J(\tau) z_{T+1-\tau} + h. \quad (39)$$

Then, we can move the temporal window of length T along the whole experimental series in order to generate new points. Then, in general, the points in the synthetic time series are generated according to the rule

$$\tilde{z}_n = \sum_{\tau=1}^T J(\tau) z_{T+n-\tau} + h. \quad (40)$$

We stress that, in our procedures, we used a stochastic version of the synthetic generation, in which the inferred vector $J(\tau)$ is replaced with its noisy version $\tilde{J}(\tau) = J(\tau) + \eta \sigma_{J(\tau)}$ with η is a $\mathcal{N}(0, 1)$ random variable. The synthetic series is then generated by collecting the $N - T$ points \tilde{z}_n .

Ethical statement. Informed consent was obtained from all subjects and related data have been treated in a completely anonymous form (in the full respect of the *Declaration of Helsinki* (1964) with the understanding and the consent of the human subjects involved in the study). APH and POLISA asked for explicit approval of the study by the responsible Ethical Committee: this approval was released to APH and POLISA on June 09 2016 by the Ethical Committee of Regione Marche (APH Hospital belongs to that region) and can be supplied upon request. All the methods were carried out in strict accordance with all the relative guidelines and regulations in Italy.

Received: 16 February 2020; Accepted: 28 July 2020

Published online: 18 September 2020

References

1. Agliari, E. et al., Detecting cardiac pathologies via machine learning on the clinical markers based on heart rate variability. *Sci. Rep.* **10**, 8845 (2020).
2. Barra, O.A. & Moretti, L., The “Life Potential” a new complex algorithm to assess “heart rate variability” from Holter records for cognitive and diagnostic aims, available at [arXiv:1310.7230](https://arxiv.org/abs/1310.7230), (2013).
3. Jaynes, E. T. *Probability theory: The logic of science* (Cambridge University Press, Cambridge, 2003).
4. Sakellariou, J., Tria, F., Loreto, V. & Pachet, F. Maximum entropy models capture melodic styles. *Sci. Rep.* **7**, 9172–9185 (2017).
5. Tkacik, G. et al. The simplest maximum entropy model for collective behavior in a neural network. *JSTAT* **2013**(03), 3011–3043 (2013).
6. Mora, T. et al. Maximum entropy models for antibody diversity. *Proc. Natl. Acad. Sci. USA* **107**(12), 5405–5410 (2010).
7. Schneidman, E. et al. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**(7087), 1007–1012 (2006).
8. Bialek, W. et al. Statistical mechanics for natural flocks of birds. *Proc. Natl. Acad. Sci. USA* **109**(13), 4786–4791 (2012).
9. Tao, W., Jin, H. & Liu, L. Object segmentation using ant colony optimization algorithm and fuzzy entropy. *Patt. Recogn. Lett.* **28**(7), 788–796 (2007).
10. Mezard, M., Parisi, G. & Virasoro, M. A. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications* (World Scientific Publishing Company, New York, 1987).
11. Shaffer, F. & Ginsberg, J. P. An overview of heart rate variability metrics and norms. *Front. Public Health* **5**, 258–287 (2017).

12. Sarlis, N. V., Skordas, E. S. & Varotos, P. A. Heart rate variability in natural time and $1/f$ noise. *Europhys. Lett.* **87**, 18003–19009 (2009).
13. Pilgram, B. & Kaplan, D. T. Nonstationarity and $1/f$ noise characteristics in heart rate. *Am. Phys. Soc.* **0363**, 6119–6125 (1999).
14. Melillo, P. *et al.* Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis. *Plos-One* **10**(3), 8504–8512 (2015).
15. Laborde, S., Mosley, E. & Thayer, J. F. Heart rate variability and cardiac vagal tone in psychophysiological research. *Front. Psychol.* **3**, 213–229 (2017).
16. Voss, A., Schulz, S., Schroeder, R., Baumert, M. & Caminal, P. Methods derived from non-linear dynamics for analysis heart rate variability. *Phil. Trans. R. Soc A* **367**, 277–296 (2009).
17. Sassi, R. *et al.* Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Hearh Rhythm Association co-endorsed by the Asia-Pacific Heart Rhythm Society. *Europace* **17**, 1341–1348 (2015).
18. McCraty, R. & Shaffer, F. Heart rate variability: New perspectives on physiological mechanisms, assesment of self-regulatory capacity and health risk. *Global Adv. Health Med.* **4**(1), 46–61 (2015).
19. Beauchine, T. P. & Thayer, J. F. Heart rate variability as a transdiagnostic biomarker of psychopathology. *Int. J. Psychophysiol.* **98**, 338–350 (2015).
20. Li, K., Rudiger, H. & Ziemssen, T. Spectra analysis of heart rate variability: Time window matters. *Front. Neurol.* **10**, 545–561 (2019).
21. Ivanov, P. C. *et al.* Scaling and universality in heart rate variability distributions. *Phys. A* **249**, 587–593 (1998).
22. Jonason, K., Vincent, E., Hammann, J., Bouchaud, J. P. & Nordblad, P. Memory and chaos effects in spin glasses. *Phys. Rev. Lett.* **81**(15), 3243–3247 (1998).
23. Banavar, J. R. & Bray, A. J. Chaos in spin glasses: A renormalization-group study. *Phys. Rev. B* **35**(16), 8888–8900 (1987).
24. Kondor, I. On chaos in spin glasses. *J. Phys. A* **22**(5), L163–L169 (1989).
25. Van Mourik, J. & Coolen, A. C. C. Cluster derivation of Parisis RSB solution for disordered systems. *J. Phys. A* **34**(10), 111–1116 (2001).
26. Weissman, M. B. $1/f$ noise and other slow nonexponential kinetics in condensed matter. *Rev. Mod. Phys.* **60**(2), 537 (1988).
27. Gongwen, P. & Herrmann, H. J. Density waves and $1/f$ density fluctuations in granular flow. *Phys. Rev. E* **51**(3), 1745 (1995).
28. Peng, C. K. *et al.* Long-range correlations in nucleotide sequences. *Nature* **356**(6365), 168–170 (1992).
29. Mercik, S., Weron, K. & Siwy, Z. Statistical analysis of ionic current fluctuations in membrane channels. *Phys. Rev. E* **60**(6), 7343–7363 (1999).
30. Bak, P., Chao, T. & Wiesenfeld, K. Self-organized criticality: An explanation of the $1/f$ noise. *Phys. Rev. Lett.* **59**(4), 381–385 (1987).
31. Christensen, K., Zeev, O. & Bak, P. Deterministic $1/f$ noise in nonconservative models of self-organized criticality. *Phys. Rev. Lett.* **68**(16), 2417–2421 (1992).
32. Ocio, M., Bouchiat, H. & Monod, P. Observation of $1/f$ magnetic fluctuations in a spin glass. *J. Phys. Lett.* **46**(14), 647–652 (1985).
33. Refregier, P. *et al.* Equilibrium magnetic fluctuations in spin glasses: Temperature dependence and deviations from $1/f$ behaviour. *Europhys. Lett.* **3**(4), 503–509 (1987).
34. Pytte, E. & Imry, Y. Ubiquity of logarithmic scaling, $1/f$ power spectrum, and the $\pi/2$ rule. *Phys. Rev. B* **35**(3), 1465–1478 (1987).
35. Baiesi, M. & Paczuski, M. Scale-free networks of earthquakes and aftershocks. *Phys. Rev. E* **69**(6), 066106–066122 (2004).
36. Rammal, R., Tannous, C. & Tremblay, A. M. S. $1/f$ noise in random resistor networks: Fractals and percolating systems. *Phys. Rev. A* **31**(4), 2662–2681 (1985).
37. Melillo, P. *et al.* Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis. *PLoS ONE* **10**(3), 8504–8510 (2015).
38. Peltola, M. A. Role of editing of R-R intervals in the analysis of heart rate variability. *Front. Physiol.* **3**, 148–166 (2012).
39. Duchi, J., Hazan, E. & Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **12**, 2121–2159 (2011).
40. Derrida, B., Gardner, E. & Zippelius, A. An exactly solvable asymmetric neural network model. *Europhys. Lett.* **4**, 167 (1987).
41. Derrida, B. & Weisbuch, G. Dynamical phase transitions in 3-dimensional spin glasses. *Europhys. Lett.* **4**, 657 (1987).
42. Crisanti, A. & Sompolinsky, H. Dynamics of spin systems with randomly asymmetric bonds: Ising spins and Glauber dynamics. *Phys. Rev. A* **37**, 4865 (1988).
43. Roudi, Y. & Hertz, J. Dynamical TAP equations for non-equilibrium Ising spin glasses. *J. Stat. Mech.* **30**, 31 (2011).
44. Campajola, C., Lillo, F. & Tantari, D. Inference of the kinetic Ising model with heterogeneous missing data. *Phys. Rev. E* **99**, 062138 (2019).

Acknowledgements

F.A. gratefully acknowledges support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, ERC Starting Grant "INTERCELLMED" (Grant Agreement No 759959). The author would like to acknowledge fruitful discussions with Dr. Loretta L. del Mercato in CNR-Nanotec, Lecce, IT. All the other Authors acknowledge the POR CALABRIA FESR/FSE 2014/2020, "MATCH-PHYTAGORAS" project financed by the Italian *Calabria Regional Authority* within the POR CALABRIA FESR/FSE 2014/2020, CUP: J48C17000250006. We are indebted with Regione Calabria because without their support these results could not have been achieved. Our thanks are due also to the APH-URCC staff (the Unit of Clinical and Cardiological Research) for the always accurate collection and supply of clinical data and for all the helpful discussions. AB acknowledges also partial financial fundings by 'Stochastic Models for Complex Systems' (PRIN 2017, Project no. 2017JFFHSH).

Author contributions

The Authors equally contributed to the research and the production of the manuscript. In particular, under the coordination of O.A.B. and L.M., data have been stored and handled by L.F.V. and O.A.B. The statistical analysis over these datasets carried by F.A., A.F., E.A. and A.B. The overall research has been coordinated by O.A.B. and L.M.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to O.A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020