HAL
open science

Information Theory to probe Intrapartum Fetal Heart Rate Dynamics

Carlos Granero-Belinchon, Stéphane G. Roux, Patrice Abry, Muriel Doret, Nicolas B. Garnier

To cite this version:

Carlos Granero-Belinchon, Stéphane G. Roux, Patrice Abry, Muriel Doret, Nicolas B. Garnier. Information Theory to probe Intrapartum Fetal Heart Rate Dynamics. Entropy, 2017, 19 (12), 10.3390/e19120640 . hal-01646840

HAL Id: hal-01646840

https://hal.science/hal-01646840v1

Submitted on 27 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Information Theory to probe Intrapartum Fetal Heart Rate Dynamics

Carlos Granero-Belinchon(1), Stéphane G. Roux(1), Patrice Abry(1),
Muriel Doret(2), Nicolas B. Garnier(1)

[1] Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, F-69342 Lyon, France;
firstname.lastname@ens-lyon.fr

[2] Femme-Mère-Enfant Hospital, Bron,, Univ. Lyon I, Lyon, France;
muriel.doret-dion@chu-lyon.fr

November 27, 2017

Abstract

Intrapartum fetal heart rate (FHR) monitoring constitutes a reference tool in clinical practice to assess the baby health status and to detect fetal acidosis. It is usually analyzed by visual inspection grounded on FIGO criteria. Characterization of Intrapartum fetal heart rate temporal dynamics remains a challenging task and continuously receives academic research efforts. Complexity measures, often implemented with tools referred to as *Approximate Entropy* (ApEn) or *Sample Entropy* (SampEn), have regularly been reported as significant features for intrapartum FHR analysis. We explore how Information Theory, and especially *auto mutual information* (AMI), is connected to ApEn and SampEn and can be used to probe FHR dynamics. Applied to a large (1404 subjects) and documented database of FHR data, collected in a French academic hospital, it is shown that i) auto mutual information outperforms ApEn and SampEn for acidosis detection in the first stage of labor and continues to yield the best performance in the second stage; ii) Shannon entropy increases as labor progresses, and is always much larger in the second stage; iii) babies suffering from fetal acidosis additionally show more structured temporal dynamics than healthy ones and that this progressive structuration can be used for early acidosis detection.

1 Introduction

Intrapartum fetal heart rate monitoring Because it is likely to provide obstetricians with significant information related to the health status of the fetus during delivery, intrapartum fetal heart rate (FHR) monitoring is a routine procedure in hospitals. Notably, it is expected to permit to detect fetal acidosis, which may induce severe consequences for both the baby and the mother and thus requires a timely and relevant decision for rapid intervention and operative delivery [1]. In daily clinical practice, FHR is mostly inspected visually, trained from clinical guidelines formalized by the International Federation of Gynecology and Obstetrics (FIGO) [2, 3]. However, it has been well documented that such visual inspection is prone to severe inter-individual variability and even shows a substantial intra-individual variability [4]. This is reflecting both that FHR temporal dynamics are complex and hard to assess, and that FIGO criteria lead to a demanding evaluation, as they mix several aspects of FHR dynamics (baseline drift, decelerations, accelerations, long and short-term variabilities). Difficulties in performing objective assessment of these criteria has led to a substantial number of unnecessary Caesarean sections [5]. This has triggered a large amounts of researches world wide aiming both to compute in a reproducible and objective way the FIGO criteria [2], and to devise new *signal processing* inspired features to characterize FHR temporal dynamics (cf. [6, 7] for reviews).

Related works After the seminal contribution in the analysis of heart rate variability (HRV) in adults [8], spectrum estimation has been amongst the first signal processing tool that has been considered for computerized analysis of FHR, either constructed on models driven by characteristic time scales [9, 10, 11] or scale-free paradigm [12, 13, 14]). Further, aiming to explore temporal dynamics beyond the mere temporal correlations, several variations of nonlinear analysis have been envisaged both for antepartum and intrapartum FHR [15], based, e.g., on multifractal analysis [13], scattering transforms [16], phase-driven synchronous pattern averages [17] or complexity and entropy measures [18, 19, 20]. Interested readers are

referred to e.g. [6, 7] for overviews. There has also been several attempts to combine features different in nature by doing multivariate classification using supervised machine learning strategies (cf. e.g., [21, 6, 22, 23, 24]).

Measures from Complexity Theory or Information Theory remain however amongst the most used tools to construct HRV characterization. They are defined independently from (deterministic) dynamical system or from (random) stochastic process frameworks. The former led to standard references, both for adult and for antepartum and intrapartum fetal heart rate analysis: *Approximate Entropy* (ApEn) [18, 25], *Sample Entropy* (SampEn) [26], which can be regarded as practical approximations to Kolmogorov-Sinai or Eckmann-Ruelle complexity measures. The stochastic process framework leads to the definitions of Shannon and Rényi entropies and entropy rates. Both worlds are connected by several relations, cf. e.g., [27, 28] for reviews. Implementations of ApEn and SampEn rely on *correlation integral* based algorithm (CI) [18, 29] while that of Shannon entropy rates may instead benefit from *k-nearest neighbor* (k-NN) algorithm [30], which brings robustness and improved performance to entropy estimation [31, 32, 33].

Labor stages Automated FHR analysis is complicated by the existence of two distinct stages during labor. The dilatation stage (stage 1) consists in progressive cervical dilatation and regular contractions. The active pushing stage (stage 2) is characterized by a fully dilated cervix and expulsive contractions. The most common approaches in FHR analysis consist in not distinguishing stages and performing a global analysis [34, 23] or in focusing on stage 1 only, as it is better documented and usually shows data with better quality, cf. e.g. [35, 24]. Whether or not temporal dynamics associated to each stage are different has not been intensively explored yet (see a contrario [36, 37]). However, recently, some contributions start to conduct systematic comparisons [38, 39].

Goals, contributions and outline The present contribution remains in the category of works aiming to design new efficient features for FHR, here based on advanced information theoretic concepts. These new tools are applied to a high quality large (1404 subjects) and documented FHR database collected across years in an academic hospital in France, and described in Section 2. The database is split into two datasets associated each with one stage of labor, which enables us first to assess and compare acidosis detection performance achieved by the proposed features independently on each stage, and second to address differences in FHR temporal dynamics between the two stages. Reexamining formally the definitions of entropy rates in Information Theory, Section 3.1 first establishes that they can be split into two components: Shannon entropy that quantifies data static properties, and Auto mutual information (AMI) that characterizes temporal dynamics in data, combining both linear and nonlinear (or higher order) statistics. ApEn and SampEn, defined from Complexity Theory, are then explicitly related to entropy rates, and hence to AMI (cf. Section 3.3). Estimation procedures for Shannon entropy, entropy rate and AMI, based on *k-nearest neighbor* (k-NN) algorithms [30], are compared to those of ApEn and SampEn, constructed on *Correlation Integral* algorithms [18, 40, 29]. Acidosis detection performances are reported in Section 4.1. Results are discussed in terms of quality versus analysis window size, k-NN or correlation integral based procedures, and differences between stage 1 and stage 2. Further, a longitudinal study consisting of sliding analysis in overlapping windows across the delivery process permits to show that processes characterizing stage 1 and stage 2 are different (Section 4.4).

2 Datasets: intrapartum fetal heart rate times series and labour stages.

Data collection. Intrapartum fetal heart rate data were collected at the academic Femme-Mère-Enfant hospital, in Lyon, France, during daily routine monitoring across the years 2000 to 2010. They are recorded using STAN S21 or S31 devices with internal scalp electrodes at 12bit resolution, 500 Hz sampling rate (STAN, Neoventa Medical, Sweden). Clinical information was provided by the obstetrician in charge, reporting delivery conditions as well as the health status of the baby, notably the umbilical artery pH after delivery and the decision for intervention due to suspected acidosis [41].

Datasets. For the present study, subjects were included using criteria detailed in [41, 24], leading to a total of 1404 tracing, lasting from 30 minutes to several hours. These criteria essentially aim to reject subjects with too low quality recording (e.g., too many missing data, too large gaps, too short recordings, ...). As a result, for subjects in the database, the average fraction of data missing in the last 20 minutes before delivery is less than 5%. The first goal of the present work is to assess the relevance of new information theoretics measures; their robustness to poor quality data is postponed for future work.

The measurement of pH, performed by blood test immediately after delivery is systematically documented and used as ground-truth: When pH ≤ 7.05, the newborn is considered has having suffered from acidosis, and referred to as acidotic (A, pH ≤ 7.05). Conversely, when pH > 7.05, the newborn is considered not having suffered from acidosis during delivery, and termed normal (N, pH > 7.05). In order to have a meaningful pH indication, we retain only subjects for which the time between end of recording and birth is lower or equal to 10 min.

Following the discussion above on labor stages, subjects are split into two different datasets. Dataset I consists of subjects for which delivery took place after a short stage 2 (less than 15 min) or during stage 1 (stage 2 was absent), It contains 913 normal and 26 acidotic subjects. Dataset II gathers FHR for delivery that took place after more than 15 min of stage 2. It contains 450 normal and 15 acidotic subjects.

Beat-per-Minute time series and preprocessing. For each subject, the collection device provides us with a digitalized list of RR-interarrivals Δ_k in ms. In reference to common practice in FHR analysis, and for ease of comparisons amongst subjects, RR-interarrivals are converted in regularly sampled Beat-per-Minute times series, by linear interpolation of the samples $\{\ldots, 36000/\Delta_k, \ldots\}$. The sampling frequency has been set to $f_s = 10$Hz as FHR do not contain any relevant information above 3Hz.

3 Methods

Outline We describe in this section the five features that we use to analyze heart rate signals. We propose to apply Information Theory, as defined by Shannon, to the analysis of cardiac signals. We do so by computing the Shannon entropy, the Shannon entropy rate, and the auto mutual information. The first section below is devoted to the definition of these quantities, which provide three features. The second section reports the definitions of two features rooted in Complexity theory: Approximate Entropy (ApEn) and Sample Entropy (SampEn), which are classics in cardiac signal analysis. Although we use them in practice only as benchmarks, we devote a last section to their relation with the new features we propose.

Information Theory and Complexity theory only differ in the nature of the objects under study. Information Theory, on one hand, aims to analyze random processes and defines functionals of probability densities. Complexity theory, on the other hand, aims to analyze signals produced by dynamical systems and assumes the existence of ergodic probability measures to describe the density of trajectories in phase space, so that they can be manipulated as probability densities. In this spirit, we consider through all this paper, the signals to analyze as random processes, although they indeed originate from a dynamical system.

Assumptions For the sake of simplicity in the description of the features, and for practical use, we assume that signals are monovariate (unidimensional) and centered (zero mean) because the five features we use are independent of the first moment of the probability density function. We also assume that signals are stationary. Although this may seem at first a very strong assumption, it is very reasonable when examining time windows smaller than the natural time scale of the evolution between stages 1 and 2, as we discuss in section 4.4, and larger than events such as contractions. Finally, we also assume that the signals contain N points, sampled at a constant frequency. All estimates depends on N via finite size effects. In the following, we do not mention this dependence explicitly in the notations, and only compare features computed over the same window size.

Time-embedding Because we are interested in the dynamics of the signal, we use the delay-embedding procedure introduced by Takens [42] in the context of dynamical systems. Its goal is to include time-correlations into the statistics. We construct, by sampling the initial process X every τ points in time, a m-dimensional time-embedded process $X^{(m,\tau)}$ as:

$$\mathbf{x}_t^{(m,\tau)} = \left(x_t,\, x_{t-\tau},\, \cdots,\, x_{t-(m-1)\tau}\right) . \tag{1}$$

In practice, we have a finite number N of points in the time-series, so there are $N - m + 1$ well-defined embedded vectors.

3.1 Information Theory features

We now briefly recall definitions from Information Theory introduced by Shannon [43]. This paradigm aims to describe processes in terms of information, and it can be applied to any experimental signals.

3.1.1 Shannon and Rényi Entropies

Shannon entropy $H(X^{(m,\tau)})$ of a m-dimensional embedded process $X^{(m,\tau)}$ is a functional of its joint probability density function $p\left(\mathbf{x}_t^{(m,\tau)}\right)$ [43]:

$$H(X^{(m,\tau)}) = -\int_{\mathbb{R}^m} p\left(\mathbf{x}_t^{(m,\tau)}\right) \log\left(p\left(\mathbf{x}_t^{(m,\tau)}\right)\right) \mathrm{d}\mathbf{x}_t^{(m,\tau)} \,. \tag{2}$$

which does not depend on t, thanks to the stationarity of the signal. Shannon entropy measures the total information contained in the process $X^{(m,\tau)}$. For embedding dimension $m = 1$, it is independent of the sampling parameter τ and we write in the following $H(X^{(1,\tau)}) = H(X)$.

Rényi q-order entropy $R_q(X^{(m,\tau)})$ is defined as another functional of the probability density [44]:

$$R_q(X^{(m,\tau)}) = \frac{1}{1-q} \log\left(\int_{\mathbb{R}^m} p^q\left(\mathbf{x}_t^{(m,\tau)}\right) d\mathbf{x}_t^{(m,\tau)}\right) \,. \tag{3}$$

When $q \to 1$, the Rényi q-order entropy converges to the Shannon entropy.

3.1.2 Entropy Rates

Shannon entropy rate is defined as:

$$h(X) = \lim_{m\to\infty} \frac{H(X^{(m,\tau)})}{m} \,, \tag{4}$$

which can be shown to be equivalent to:

$$h(X) = \lim_{m\to\infty} \left(H(X^{(m+1,\tau)} - H(X^{(m,\tau)})\right) \,. \tag{5}$$

We define the m-order Shannon entropy rate as measuring how much the Shannon entropy increases between two consecutive time-embedding dimensions m and $m+1$:

$$h^{(m,\tau)}(X) = H(X^{(m+1,\tau)}) - H(X^{(m,\tau)}) \tag{6}$$

$$= H(X_{t+\tau}|X_t^{(m,\tau)}) \,. \tag{7}$$

It quantifies the variation of the total information in the time-embedded process when the embedding dimension m is increased by 1. Interpreting eq.(6), it measures information in the $m+1$ coordinates of $X_t^{(m+1,\tau)}$ that is not contained in the m coordinates of $X_t^{(m,\tau)}$. Following eq.(7), it can also be interpreted as the new information brought by the extra sample $X_{t+\tau}$ when the set of m samples, $X_t^{(m,\tau)}$, is already known.

Rényi order-q entropy rate and m-order Rényi order-q entropy rate can be defined in the same way, replacing Shannon entropy H by Rényi order-q entropy R_q in eq.(4) and eq.(6), respectively. Nevertheless, it should be emphasized that Rényi order-q entropy is lacking chain rule of conditional probabilities as soon as $q \neq 1$; therefore, eq.(7) does not hold for Rényi order-q entropy, unless $q = 1$ (Shannon entropy).

3.1.3 Mutual Information

The Mutual Information (MI) of two processes measures the information they share [43]. MI is the Kullback-Leibler divergence [45] between the joint probability function and the product of the marginals which would express the joint probability function if the two processes were independent. For time-embedded processes $X^{(m,\tau)}$ and $Y^{(p,\tau)}$, it reads:

$$I(X^{(m,\tau)}, Y^{(p,\tau)}) = I(\mathbf{x}_{t'}^{(m,\tau)}, \mathbf{y}_t^{(p,\tau)}) =$$

$$= \int_{\mathbb{R}^{m+p}} p(\mathbf{x}_{t'}^{(m,\tau)}, \mathbf{y}_t^{(p,\tau)}) \log\left(\frac{p(\mathbf{x}_{t'}^{(m,\tau)}, \mathbf{y}_t^{(p,\tau)})}{p(\mathbf{x}_{t'}^{(m,\tau)})p(\mathbf{y}_t^{(p,\tau)})}\right) \mathrm{d}\mathbf{x}_{t'}^{(m,\tau)} \mathrm{d}\mathbf{y}_t^{(p,\tau)} \,. \tag{8}$$

Mutual information is symmetrical with respect to its two arguments. If X and Y are stationary processes, the mutual information $I(X^{(m,\tau)}, Y^{(p,\tau)})$ depends only on $t' - t$, the time difference between $X^{(m,\tau)}$ and $Y^{(p,\tau)}$.

Auto Mutual Information If $Y = X$ and $t' - t = p\tau$, the MI measures the information shared by two consecutive chunks of the same process X, both sampled at τ. This quantity is sometimes called "information storage" [46, 47, 33], and we refer to it as the Auto Mutual Information (AMI) of the process X:

$$I^{(m,p,\tau)}(X) = I(\mathbf{x}_{t+p\tau}^{(p,\tau)}, \mathbf{x}_t^{(m,\tau)}) \,. \tag{9}$$

Remarking that the concatenation of $\mathbf{x}_{t+p\tau}^{(p,\tau)}$ and $\mathbf{x}_t^{(m,\tau)}$ is nothing but the $(m+p)$-dimensional embedded vector

$$\left(\mathbf{x}_{t+p\tau}^{(p,\tau)}, \mathbf{x}_t^{(m,\tau)}\right) = \mathbf{x}_t^{(m+p,\tau)} \,,$$

the AMI depends on the embedding dimensions (m, p) and the sampling time τ only. AMI of order (m, p) measures the shared information between consecutive m-points and p-points dynamics, *i.e.*, by how much the uncertainty of future p-points dynamics is reduced if the previous m-points dynamics is known.

Thanks to the symmetry of the MI with respect to its two arguments, and invoking the stationarity of X, the AMI is invariant when exchanging m and p:

$$I^{(m,p,\tau)}(X) = I(\mathbf{x}_{t+p\tau}^{(p,\tau)}, \mathbf{x}_t^{(m,\tau)}) = I(\mathbf{x}_{t+m\tau}^{(m,\tau)}, \mathbf{x}_t^{(p,\tau)}) = I^{(p,m,\tau)}(X) \,. \tag{10}$$

We emphasize that the MI — and therefore AMI — are defined only for the Shannon entropy. The expression of the Rényi order-q mutual information is not unique as soon as $q \neq 1$, and we do not consider it here.

Special case $p = 1$ If $p = 1$, the AMI is directly related to the Shannon entropy rate of order m:

$$h^{(m,\tau)}(X) = H(X) - I^{(m,1,\tau)}(X) \,, \tag{11}$$

or equivalently

$$I^{(m,1,\tau)}(X) = H(X) - h^{(m,\tau)}(X) \,. \tag{12}$$

Interestingly, this splits the entropy rate $h^{(m,\tau)}(X)$ in two contributions. The first one is the total entropy $H(X)$ of the process which only depends on the one-point statistics, and so does not describe the dynamics of X. The second term is the AMI $I^{(m,1,\tau)}(X)$ which depends on the dynamics of the process X, irrespective of its variance [48].

Special case of a process with Gaussian distribution For illustration, if X is a stationary Gaussian process, hence fully defined by its variance σ^2 and normalized correlation function $c(\tau)$, we have:

$$H^{(p,\tau)}(X) = \frac{p}{2}\log\left(2\pi e\sigma^2\right) + \frac{1}{2}\log\left(|\Sigma^{(p)}|\right) \tag{13}$$

$$I^{(m,p,\tau)}(X) = \frac{1}{2}\log\left(\frac{|\Sigma^{(m)}||\Sigma^{(p)}|}{|\Sigma^{(m+p)}|}\right), \tag{14}$$

where $\Sigma^{(m)}$ is the $m \times m$ correlation matrix of the process X; $\Sigma_{i,j} = c(|i-j|\tau)$ and $|\Sigma^{(1)}| = 1$. For the particular case $p = 1$, we have:

$$H(X) = \frac{1}{2}\log\left(2\pi e\sigma^2\right) \tag{15}$$

$$I^{(m,1,\tau)}(X) = \frac{1}{2}\log\left(\frac{|\Sigma^{(m)}|}{|\Sigma^{(m+1)}|}\right) \tag{16}$$

$$h^{(m,\tau)}(X) = \frac{1}{2}\log\left(2\pi e\sigma^2\right) - \frac{1}{2}\log\left(\frac{|\Sigma^{(m)}|}{|\Sigma^{(m+1)}|}\right), \tag{17}$$

which clearly illustrates the decomposition of the entropy rate according to eq.(11): the first term $H(X)$ depends only on the static (one-point) statistics (via σ^2) and the second term $I^{(m,1,\tau)}(X)$ depends on the temporal dynamics (and in this simple case only on the dynamics, via the auto-correlation function $c(\tau)$).

3.2 Features from Complexity Theory

In the 1960's, Kolmogorov and Sinai adapted Shannon's Information Theory to the study of dynamical systems. The divergence of trajectories starting from different but undistinguishable initial conditions can be pictured as creating uncertainty, so creating information. Kolmogorov Complexity (KC), also known as the Kolmogorov-Sinai entropy and noted $h_{\mathrm{KS}}(\rho)$ in the following, measures the mean rate of creation of information by a dynamical system with ergodic probability measure ρ. KC is constructed exactly as the

Shannon entropy rate from Information Theory, using eq.(4) and the same functional form as in eq.(2), but using the density ρ of trajectories in phase space instead of the probability density p. In the early 1980's, the Eckmann-Ruelle entropy $K_2(\rho)$ [29, 49] was introduced following the same steps but using the functional form of the Rényi order-2 entropy (equation (3)). The interest of K_2 relies in its easier and hence faster computation from experimental time series, which was at the time a challenging issue.

Kolmogorov-Sinai and Eckmann-Ruelle entropies The ergodic theory of chaos provides a powerful framework to estimate the density of trajectories in the phase space of a chaotic dynamical system [49]. For an experimental or numerical signal, it amounts to assimilating the phase space average to the time average. Given a distance $d(.,.)$ — usually defined with the $\mathcal{L}^2$ or the $\mathcal{L}^\infty$ norm — in the m-dimensional embedded space, the local quantity

$$C_i^m(\epsilon) = \frac{\text{number of } j \text{ such that } d\left(\mathbf{x}_i^{(m,\tau)}, \mathbf{x}_j^{(m,\tau)}\right) \leq \epsilon}{N-m+1}, \tag{18}$$

provides, up to a factor ϵ, an estimate of the local density ρ in the m-dimensional phase space around the point $\mathbf{x}_i^{(m,\tau)}$. The following averages:

$$\Phi^m(\epsilon) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(\epsilon), \tag{19}$$

$$C^m(\epsilon) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} C_i^m(\epsilon), \tag{20}$$

are then used to provide the following equivalent definitions of the complexity measures [49]:

$$h_{\mathrm{KS}}(\rho) = \lim_{\epsilon\to 0} \lim_{m\to\infty} \lim_{N\to\infty} \left(\Phi^m(\epsilon) - \Phi^{m+1}(\epsilon)\right), \tag{21}$$

$$K_2(\rho) = \lim_{\epsilon\to 0} \lim_{m\to\infty} \lim_{N\to\infty} \ln\left(\frac{C^m(\epsilon)}{C^{m+1}(\epsilon)}\right). \tag{22}$$

3.2.1 Approximate Entropy

Approximate Entropy (ApEn) was introduced by Pincus in 1991 for the analysis of babies heart rate [50]. It is obtained by relaxing the definition (21) of h_{KS} and working with a fixed embedding dimension m and a fixed box size ϵ, often expressed in units of the standard deviation σ of the signal as $\epsilon = r\sigma$. ApEn is defined as

$$\mathrm{ApEn}(m,\epsilon) = \Phi^m(\epsilon) - \Phi^{m+1}(\epsilon). \tag{23}$$

On the practical side, and in order to have a well-defined $\Phi^m(\epsilon)$ in (19), the counting of neighbors in the definition (18) allows self-matches $j = i$. This ensures that $C_i^m(\epsilon) > 0$, which is required by (19). ApEn depends on the number of points N in the time series. Assuming N is large enough, we have

$$\lim_{\epsilon\to 0} \lim_{m\to\infty} \mathrm{ApEn}(m,\epsilon) = h_{\mathrm{KS}}, \tag{24}$$

We interpret ApEn as an estimate of the m-order Kolmogorov-Sinai entropy h_{KS} at finite resolution ϵ. The larger N, the better the estimate. More interesting is that the non-vanishing value of ϵ in its definition makes ApEn insensitive to details at scales lower than ϵ. On one hand, this is very interesting when considering an experimental (therefore noisy) signal: choosing ϵ larger than the rms of the noise (if known) filters the noise, and ApEn is then expected to measure only the complexity of the underlying dynamics. This was the main motivation of Pincus and explains the success of ApEn. On the other hand, not taking the limits $\epsilon \to 0$ and $m \to \infty$ makes ApEn an ill-defined quantity that has no reason to behave like h_{KS}. In addition, only very few analytical results have been reported on the bias and the variance of ApEn.

Although m should in theory be larger than the dimension of the underlying dynamical system, ApEN is defined and used for any possible value of m and most applications reported in the literature are using small m (1 or 2) without any analytical support, but with great success [50, 51].

3.2.2 Sample Entropy

A decade after Pincus seminal paper, Richman and Moorman pointed out that ApEn contains in its very definition a source of bias and was lacking in some cases "relative consistency". They defined Sample Entropy

(SampEn) on the same grounds as ApEn:

$$\mathrm{SampEn}(m,\epsilon) = \ln\left(\frac{C^m(\epsilon)}{C^{m+1}(\epsilon)}\right) . \tag{25}$$

so that

$$\lim_{\epsilon\to 0}\lim_{m\to\infty}\mathrm{SampEn} = K_2 , \tag{26}$$

On the practical side, the counting of neighbors in (18) does not allow self-matches. $C_i^m(\epsilon)$ may vanish, but when averaging over all points in eq.(20), the correlation integral $C^m(\epsilon) > 0$. In practice, SampEn is considered to improve on ApEn as it shows lower sensitivity to parameter tuning and data sample size than ApEn [52, 53].

We interpret SampEn as an estimate of the m-order Eckmann-Ruelle entropy K_2 at finite resolution ϵ.

3.2.3 Estimation

We note in the following $\mathrm{ApEn}^{(m)}$ and $\mathrm{SampEn}^{(m)}$ the estimated values of ApEn and SampEn using our own Matlab implementation, based on Physio-Net packages. We used the commonly accepted value, $\epsilon = 0.2\sigma$, with σ the standard deviation of X, and $m = 2$. For all quantities, we used $\tau = 5 = f_s/f_{\max}$ with $f_{\max} = 2\mathrm{Hz}$ the cutoff frequency above which FHR times series essentially contain no relevant information [13]; this time delay corresponds to 0.5s.

3.3 Connecting Complexity Theory and Information Theory.

We consider here for clarity only the relation between ApEn and m-order Shannon entropy rate, although the very same relation holds between SampEn and m-order Rényi order-2 entropy rate. In Information theory terms, ApEn appears as a particular estimator of the m-order Shannon entropy rate that computes the probability density by counting, in the m-dimensional embedded space, the number of neighbors in a hypersphere of radius ϵ, which can be interpreted as a particular kernel estimation of the probability density.

3.3.1 Limit of large datasets and vanishing ϵ: exact relation

When the size ϵ of the spheres tends to 0, the expected value of ApEn for a stochastic signal X with any smooth probability density is related, in the limit $N \to \infty$, to the m-order Shannon entropy rate [28]:

$$\mathrm{ApEn}(m,\epsilon) \underset{\epsilon\to 0}{\simeq} h^{(m,\tau)}(X) - \log(2\epsilon) . \tag{27}$$

Both terms involve m-points correlations of the process X. This relation allows a quantitative comparison of ApEn with the m-order Shannon entropy rate $h^{(m,\tau)}$. The $\log(2\epsilon)$ difference corresponds to the paving — with hyperspheres of radius ϵ — of the continuous m-dimensional space over which the probability $p(\mathbf{x}_t^{(m,\tau)})$ involved in eq.(2) is defined, and thus $h^{(m,\tau)}$. This paving defines a discrete phase space, over which eqs. (18), (19) and (23) operate to define ApEn [54]. This illustrates that, for a stochastic signal, ApEn diverges logarithmically as the size ϵ approaches 0, as expected for h_{KS}. Fortunately, ϵ is fixed in the definition of ApEn which allows in practice to compute it for any signal/process.

3.3.2 New features

Once recognized the success of ApEn and remembering its relation to $h^{(m,\tau)}$, it seems interesting to probe other m-order Shannon entropy rate estimators. A straightforward improvement would be to consider a smooth — e.g. Gaussian — kernel of width ϵ instead of the step function used in (18). We prefer to reverse the perspective and use a k-nearest-neighbor (k-NN) estimate. Instead of counting the number of neighbors in a sphere of size ϵ, this approach searches for the size of the sphere that accomodates k neighbors. In practice, we compute the entropy H with the Kozachenko & Leonenko estimator [30, 55], which we note $\hat{H}$. We compute the auto mutual information $I^{(m,p,\tau)}$ with the Kraskov *et al* estimator [56], which we note $\hat{I}^{(m,p)}$. We then combine the two according to eq.(11) to get an estimator $\hat{h}^{(m)}$ of the m-order Shannon entropy rate. We use $k = 5$ neighbors and set $\tau = 5$ (see section 3.2.3).

We report in the next section our results for the five features when setting $m = 2, p = 1$, and compare their performances in detecting acidosis. The dependance of the m-order entropy rate (and its estimators) on m is expected to give some insight in the dimension of the attractor of the underlying dynamical system, but as we have pointed out, the dynamics is indeeed contained in the AMI part of the entropy rate. This is why we further explore the effect of varying the embedding dimensions m and p on the AMI estimator $\hat{I}^{(m,p)}$.

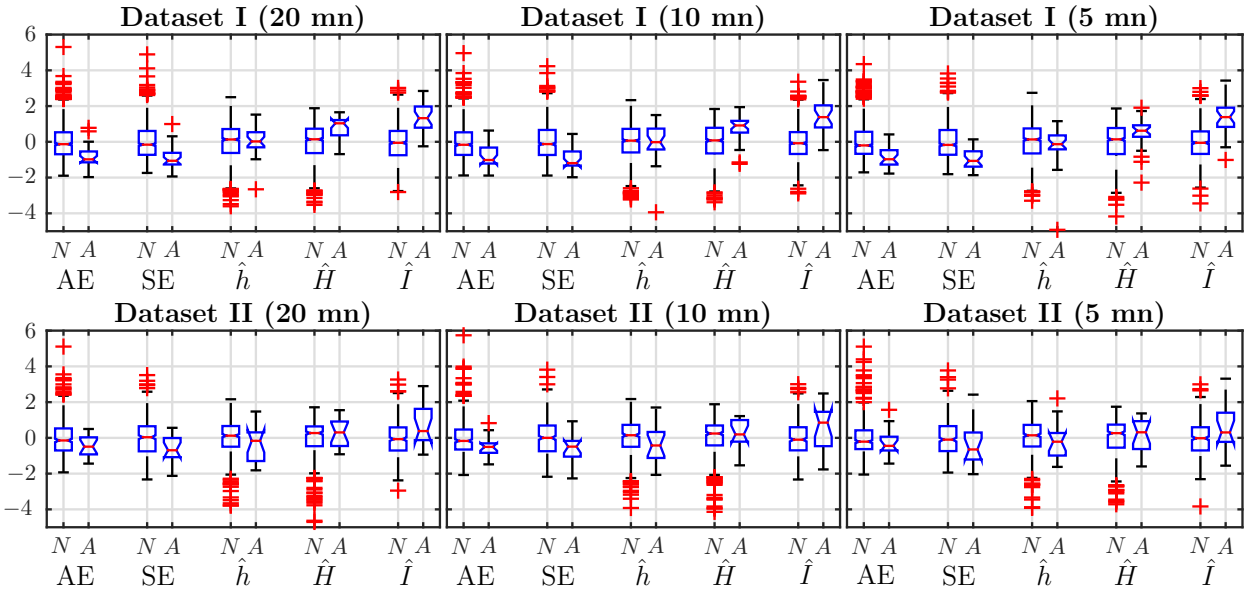

Figure 1: Box plot based comparisons of the five different (normalized) estimates ($\text{ApEn}^{(m)}$ (AE), $\text{SampEn}^{(m)}$ (SE), Shannon entropy rate $h^{(m)}$ ($\hat{h}$), Shannon entropy ($\hat{H}$) and AMI $I^{(m,p)}$ ($\hat{I}$)) for normal (N) and pathological (A for "abnormal") subjects, for dataset I (top) and dataset II (bottom). Each column corresponds to a different window size: $T = 20$mn, 10mn and 5mn. All features are computed with $m = 2$, $p = 1$.

4 Results: Acidosis detection performance

4.1 Comparison of features performance, using a single time window, just before delivery

Average features value for normal and abnormal subjects We compute the five features: $\text{ApEn}^{(m)}$, $\text{SampEn}^{(m)}$, $\hat{h}^{(m)}$, $\hat{H}$ and $\hat{I}^{(m,p)}$ for normal and acidotic (abnormal) subjects in dataset I and II using data from the last $T = 20$mn before delivery, which are the most crucial. We use the classical values $m = 2$ and $p = 1$ for embedding dimensions. To compare performance, we present the box plots of the five normalized (zero-mean, unit-variance) estimates in the left column of Fig. 1. For dataset I, the average of ApEn and SampEn for acidotic subjects is smaller than for normal subjects, while the average of Shanon entropy rate does not show any tendency. This is surprising as one might have expected for $\hat{h}^{(m)}$ a behavior similar to ApEn and SampEn (see section 3.3). Average values of $\hat{H}$ and $\hat{I}^{(2,1)}$ are larger for acidotic subjects. The larger value of Shannon entropy H indicates that the acidotic FHR signals contain more information. The larger value of AMI indicates a stronger dependence structure in the dynamics of abnormal subjects.

For subjects in dataset II, it is harder to find any tendency by looking the average values.

Features performance Fetal acidosis detection performance is assessed with the p-value given by the classical Wilcoxon ranksum test. This non-parametric test of the null hypothesis — which corresponds to identical medians of the distributions of estimates in the normal and abnormal classes — is reported in table 1. We have added one $\star$ symbol when the p-value is less than 0.05, two $\star\star$ when less than 0.01. We see that for dataset I, $\text{ApEn}^{(m)}$, $\text{SampEn}^{(m)}$, $\hat{H}$ and $\hat{I}^{(m,p)}$ for $m = 2$ discriminate normal and acidotic subjects, while $\hat{h}^{(m)}$ does not. Out of the three estimates ($\text{ApEn}^{(m)}$, $\text{SampEn}^{(m)}$, $\hat{h}^{(m)}$) based on entropy rates, the nearest-neighbors one for Shannon entropy rate is the poorest, although its decomposition into Shannon entropy (static one-point information) and AMI (which includes dynamic information) leads to two satisfying estimates Fig. 1 and Table 1 both show that the best performing estimators are $\text{SampEn}^{(m)}$ and $\hat{I}^{(m,p)}$. In dataset II, although all features performs more poorly than in dataset I, SampEn and AMI are again the best ones, with a p-value lower than 0.05. We focus on these two features in the following.

	Dataset I		Dataset II	
	AUC	p-value	AUC	p-value
ApEn$^{(2)}$	0.76	4.08e-06 $\star\star$	0.61	1.33e-01
SampEn$^{(2)}$	0.79	5.92e-07 $\star\star$	0.67	2.35e-02 $\star$
$\hat{h}^{(2)}$	0.50	9.75e-01	0.39	1.36e-01
$\hat{H}$	0.76	8.36e-06 $\star\star$	0.56	4.23e-01
$\hat{I}^{(2,1)}$	**0.84**	2.00e-09 $\star\star$	**0.68**	1.69e-02 $\star$

Table 1: Area under Receiver Operational Characteristics curves (as shown in Fig. 2) and p-value obtained from the Wilcoxon rank-sum test, for each of the five different estimates, all with $m = 2$.

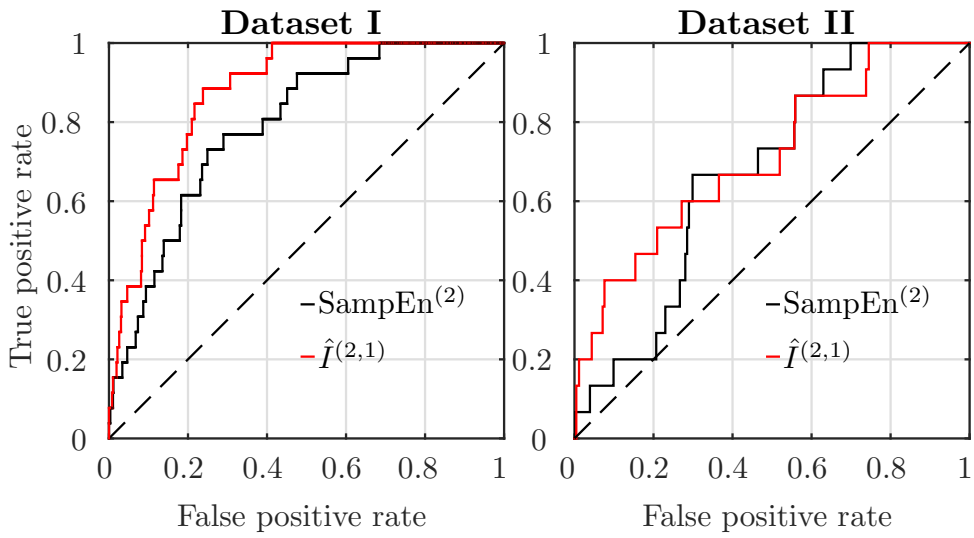

Figure 2: Receiver Operational Characteristics (ROC) curves for SampEn$^{(m)}$ (black) and AMI estimator $\hat{I}^{(m,p)}$ (red), for subjects in dataset I (left) and II (right). $m = 2$, $p = 1$.

Receiver Operational Characteristics To compare the two best performing features, SampEn and AMI, we plot receiver operational characteristics (ROC) curves in Fig. 2, both for dataset I and II, using data from the last $T = 20$mn before delivery. For dataset I, AMI better discriminates acidotic subjects from normal ones. For dataset II, AMI discrimination is only slightly better than SampEn. The area under the curve (AUC) of the ROC of the features is reported in Table 1, with a bold font to indicate the estimator with the largest AUC. Performance is much worse in dataset II than in dataset I: the AUC is reduced. Nevertheless, AMI is always the better performing estimator (its AUC reduces from 0.84 to 0.68), followed by SampEn (AUC reducing from 0.79 to 0.67).

4.2 Effect of the window size on the performance

We investigate the robustness of the detection performance when the window size T is reduced, using data from $T = 20, 10$ and 5 minutes. Results are reported in Fig 1 and Table 2. p-values and AUC both indicate that $\hat{I}^{(2,1)}$ and SampEn$^{(2)}$ provide robust discrimination in dataset I even when the observation length is reduced. Again, $\hat{I}^{(2,1)}$ is better performing: its AUC is reduced from 0.84 to 0.83 when T is reduced from 20mn to 5mn, where the AUC of SampEn is reduced from 0.79 to 0.72. In dataset II, once again, performance degrades, but AMI is still better at discriminating acidotic from normal subjects. In the following, we focus on AMI estimates only.

		SampEn$^{(2)}$		$\hat{I}^{(2,1)}$		$\hat{I}^{(3,3)}$	
		AUC	p-value	AUC	p-value	AUC	p-value
Dataset I	20 mn	0.79	5.92e-07 $\star\star$	0.84	2.00e-09 $\star\star$	**0.88**	5.46e-11 $\star\star$
	10 mn	0.76	1.22e-07 $\star\star$	0.84	2.22e-09 $\star\star$	**0.87**	1.40e-10 $\star\star$
	5 mn	0.72	1.97e-07 $\star\star$	0.83	7.47e-09 $\star\star$	**0.86**	6.26e-10 $\star\star$
Dataset II	20 mn	0.67	2.35e-02 $\star$	0.68	1.69e-02 $\star$	**0.71**	5.36e-03 $\star\star$
	10 mn	0.62	1.56e-02 $\star$	0.64	5.87e-02	**0.68**	1.66e-02 $\star$
	5 mn	0.62	5.16e-02	0.60	1.70e-01	**0.64**	7.29e-02

Table 2: AUC and p-value of Wilcoxon test of SampEn and AMI in datasets I and II using data from the last 20, 10 or 5mn before delivery.

4.3 Effect of the embedding dimensions on (fetal acidosis detection) performance of AMI

In order to improve the acidosis detection performance of the AMI, especially in dataset II, we increase the embedding dimensions m and p used in computing $\hat{I}^{(m,p)}$. This way, we probe higher order dependence stucture in the dynamics. Because of the symmetry of AMI (eq. (10)), and aiming at probing the effect of increasing either m or p, we plot the AUC of ROC as a function of $m+p$ only. The dependence of AMI on $m-p$ is much smaller and not reported here. These computations have been done with a larger value $k=15$ in the k-NN algorithms, in order to accommodate the possibly large embedding dimensions ($m+p$ up to 12). Results are presented in Fig. 3.

For a fixed window size, the AUC increases when $m+p$ increases, and reaches a maximum; it then remains constant or decreases slightly. This behaviour is observed in both datasets I and II and for any window size $T \in [5, 10, 20]$mn. Varying T dos not seems to change the location of the maximum of the AUC in a given dataset. The optimal embedding dimension is $m+p=6$ in dataset I and $m+p=10$ in dataset II. This hints at a difference in the dynamics of the FHR in the two datasets. Because both bias and computation time increase with the total dimensionality [57], the maximal embedding is restricted to $m=p=3$. A reduction of the AUC is observed when the analysis window is reduced, but this is only significant for dataset II.

We reported in Table 2 the AUC and p-values of the AMI for two embedding dimensions: $\hat{I}^{(2,1)}$ and $\hat{I}^{(3,3)}$, for datasets I and II and several window sizes. The best performing estimator is indicated in bold. For all observation windows and for the two datasets, $\hat{I}^{(3,3)}$ achieves the best performance. Their AUC is always larger than the one obtained using SampEn or AMI with $m=2$ and $p=1$.

4.4 Dynamical Analysis

We now explore how long before delivery can the AMI diagnose fetal acidosis on a FHR signal. To do so, we do not restrict our analysis to the last data points before delivery, but we apply it to an ensemble of windows scanning first and second stages of labor. We examine the dynamics of $\hat{I}^{(3,3)}$, the best performing feature, for both normal and abnormal subjects.

4.4.1 Dataset I: rapid delivery

In this first section, we focus on dataset I and probe the stage 1, including early labor, active labor, and transition. Using the time at which stage 1 ends as a reference (setting it at $t_0=0$), we compute for each subject $\hat{I}^{(3,3)}$ in a set of time windows $[t_i - T, t_i]$, $0 \leq i \leq 50$ of fixed size T ending at $t_i = t_0 - i * 2$mn, so separated by 2mn. We perform this analysis for three window sizes $T \in [20, 10, 5]$mn. The value of AMI computed in the i-th window is then assigned to time $t_i - T/2$, at the center of the interval. By construction of dataset I, delivery occurs less than 15 minutes after pushing started, and can be as short as 1mn, so we completely discard data from stage 2. We then average the values of AMI over the population of normal subjects, and over the population of acidotic subjects, respectively. Results, including p-values, are presented in Fig.4.

A first observation is that AMI is always larger for acidotic subjects than for normal subjects. As labor progresses, AMI increases in both populations, but the increase is stronger for acidotic subjects. As a consequence, the p-value of the test decreases clearly, so the feature performs better and better when approaching delivery. Detection of acidosis using the AMI feature and $T=20$mn can be obtained in dataset

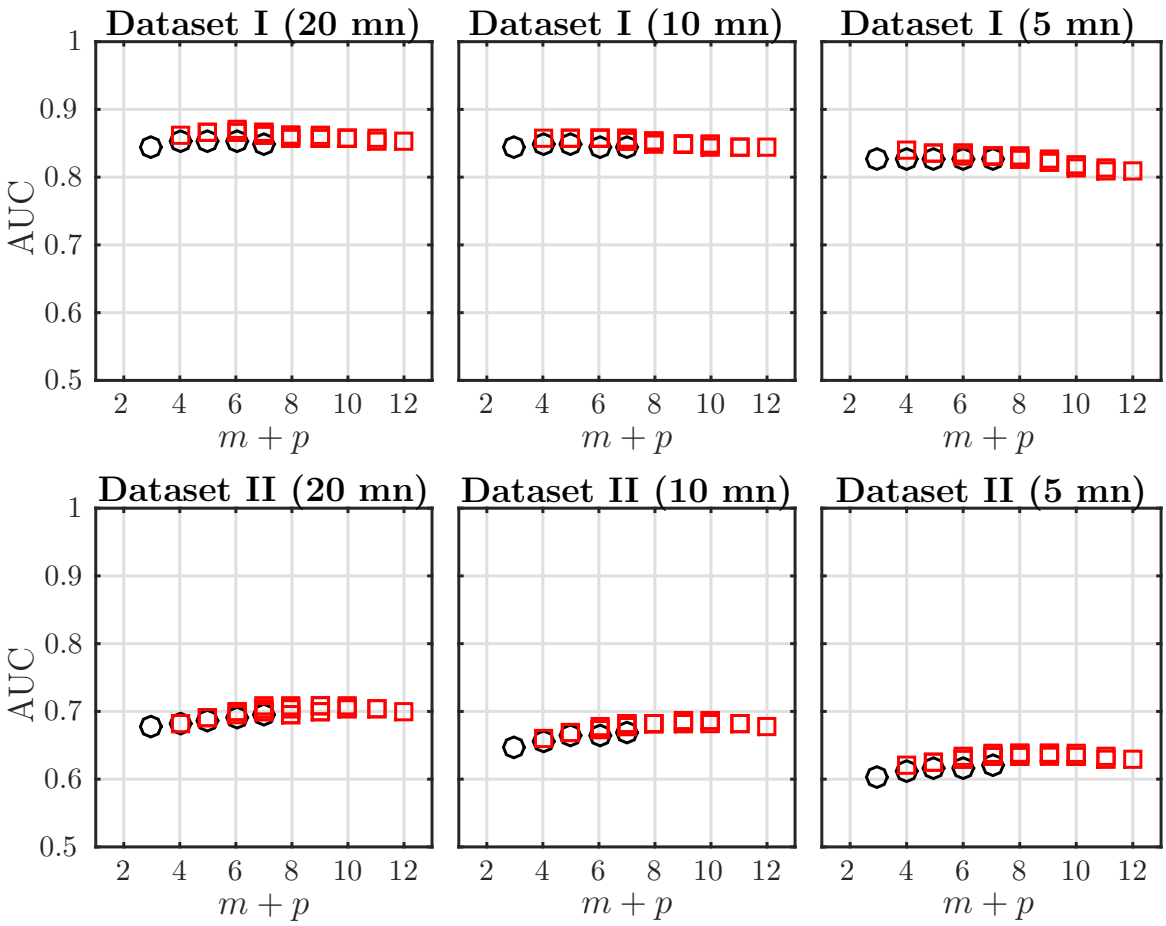

Figure 3: AUC of ROC for $\hat{I}^{(m,p)}$ as a function of the total embedding dimension $m+p$, with $m \geq p$, for time windows of size $T = 20, 10, 5$mn, for data in dataset I (first line) and dataset II (second line). Black circles indicate the special case $p = 1$ corresponding to the classical definition of Shannon entropy rate (see eq. (11)). Red squares correspond to $p \geq 2$.

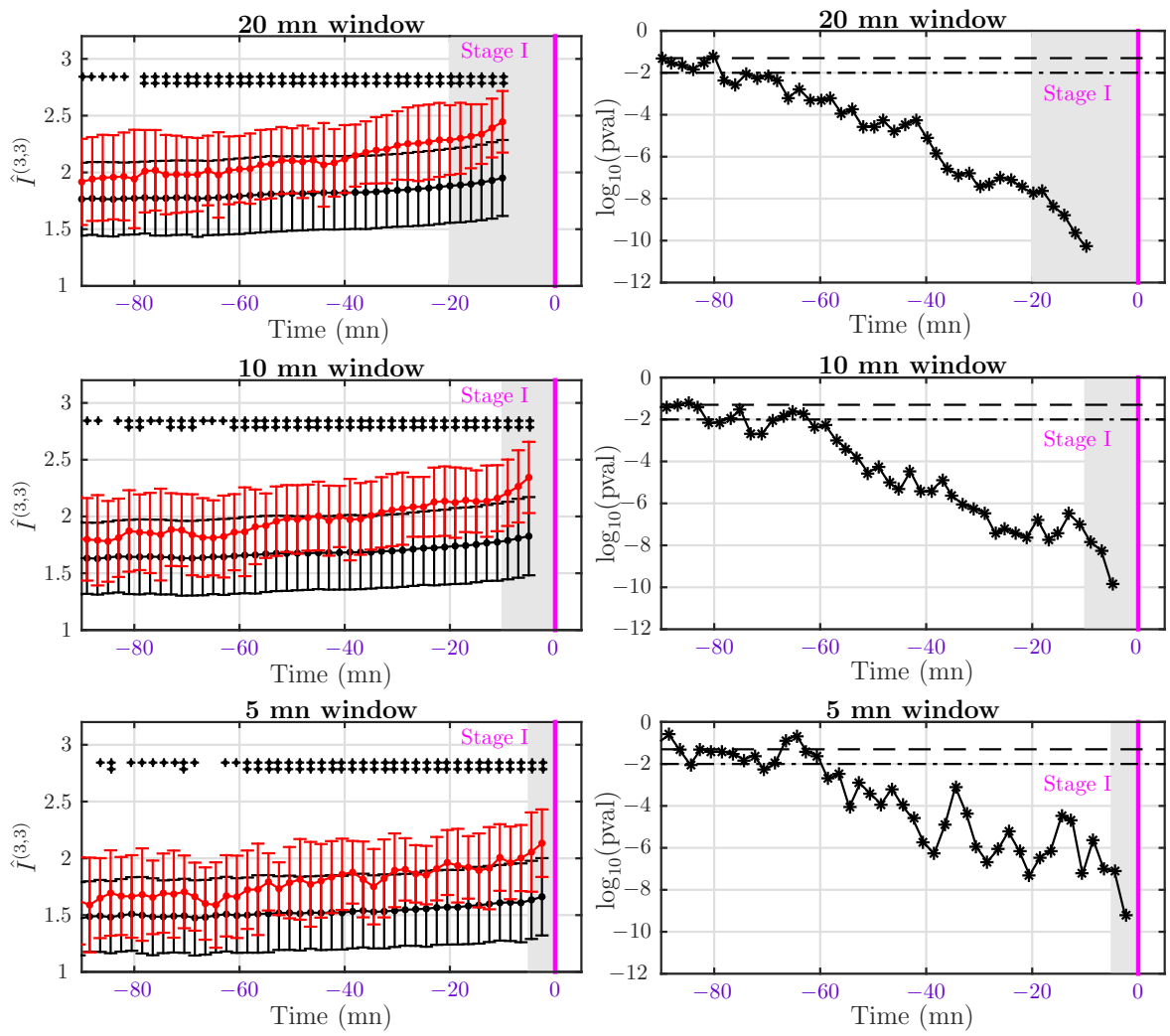

Figure 4: Left: average AMI $\hat{I}^{(3,3)}$ for normal (black) and abnormal (red) subjects in dataset I (delivery occuring less then 15mn after pushing started). AMI is computed in windows of size $T = 20$mn (first line), 10mn (second line) and 5mn (third line) shifted by 2mn. Vertical magenta line indicate the beginning of stage 2 (pushing). Right: corresponding p-value. A single black + symbol in the AMI plot indicates a p-value lower than 0.05, two ++ indicate a p-value lower than 0.01. The gray-shaded region represents the time window used to compute the last value of AMI.

I as early as 80 minutes before entering in second stage. Using shorter windows, $T = 10$mn or 5mn, detection is still reliable as early as one hour before stage 2. We interpret this reduced forecast of acidosis detection in dataset I as a direct consequence of the reduction of the statistics when the window size T is reduced.

4.4.2 Dataset II : delivery after pushing more than 15mn

For dataset II, we performed the same dynamical analysis as in the previous section, using the end of stage 1 as the reference time ($t_0 = 0$). Because there is now enough data in the pushing stage, we also perform the analysis of this stage using the delivery time ($t_0 = D$) as the reference. All results are presented in Fig. 5. At the end of stage 1, we observe again that AMI is larger for acidodic subjects than for normal ones, but the difference is not significant in this group (see the corresponding p-value on the right of Fig. 5). The situation is identical at the end of stage 2, although we obtain a lower p-value in some windows. The p-value does not decrease clearly when approaching delivery time, as it was in dataset I, see Fig. 4. For subjects in dataset II it is very difficult to make an early detection of acidosis. However, we observe in Fig. 5 that the average AMI is significantly larger at the end of stage 2 than at the end of stage 1. The increase of AMI is larger for abnormal subjects.

To examine more precisely the dynamical increase of AMI, especially when entering stage 2, we computed $\hat{I}^{(3,3)}$ over an ensemble of windows of size $T = 20$mn spanning continuously a large time interval that includes the end of the active labor and the beginning of the pushing stage. Results are reported in Fig. 6. We see a continuous increase of AMI values when evolving from stage 1 to stage 2. The increase is more important for abnormal subjects which corroborates the findings in Fig. 5. For smaller window sizes, the situation is less clear.

We also studied the dynamical evolution of the Shanon entropy estimate $\hat{H}$, which, together with the AMI, combines into the Shannon entropy rate (see eq.(11)). Changes in the Shannon entropy H indicate changes in the probability density of the signal. Results are presented in Fig. 6, side by side with the AMI. We observe a dramatic rise of the value of $\hat{H}$ when subjects evolve from stage 1 to stage 2. This increase is clearly observed for normal and abnormal subjects. No significant difference between normal and acidotic subjects is observed for this static quantity. The start of pushing implies a strong deformation of the probability density of the FHR, indicating strong perturbations of the FHR, for both normal and acidotic subjects.

5 Discussion, conclusions and perspectives

We now discuss interpretation of Shannon entropy and AMI measurements in different stages of labour. The fetuses are classified as normal or acidotic depending on a post-delivery pH measurement, which gives a diagnosis of acidosis at delivery only. There is no information on the health of the fetuses during labor.

The physiological interpretation of a feature, and especially its relation to specific FHR patterns, e.g., like those detailed in [58, 2], is a difficult task that is only scarcely reported in the litterature [59, 60]. In this article, we have averaged our results over large numbers of (normal or acidodic) subjects, which jeopardizes any precise interpretation in terms of a specific FHR pattern that may appear only intermittently.

5.1 Acidosis detection in first stage

We can nevertheless suggest that the value of the Shannon entropy H is related to the frequency of decelerations in the FHR signals. Indeed, Shannon entropy strongly depends on the standard deviation of the signal (e.g., see eq.(15)), which in turn depends on the variability in the observation window. A larger number of decelerations in the observation window deforms the PDF of the FHR signal by increasing its lower tail; in particular, this increases the width of the PDF, and hence increases the standard deviation and the Shannon entropy. This explains our findings in Fig. 1 (for dataset I).

When acidosis develops in first stage of labor, the Shannon auto mutual information estimator $\hat{I}^{(m,p)}$ significantly outperforms all other quantities both in terms of p-value and AUC. The performance of AMI is robust when tuning either the size of the observation — and hence the number of points in the data — and the embedding dimensions (m, p). In addition, the performance slightly increases when the total embedding dimension $m + p$ increases; although one has to care about the curse of dimensionality.

For abnormal subjects from dataset II, AMI is not able to detect acidosis using data from stage 1. This suggests that acidosis develops later, in the second stage of labor.

For all datasets, AMI computed with $\tau = 0.5$s is always larger for acidotic subjects than for normal subjects. This is in agreement with results obtained with ApEn and SampEn, which are both lower for acidotic subjects. This shows that FHR classified as abnormal have a stronger dependence structure at

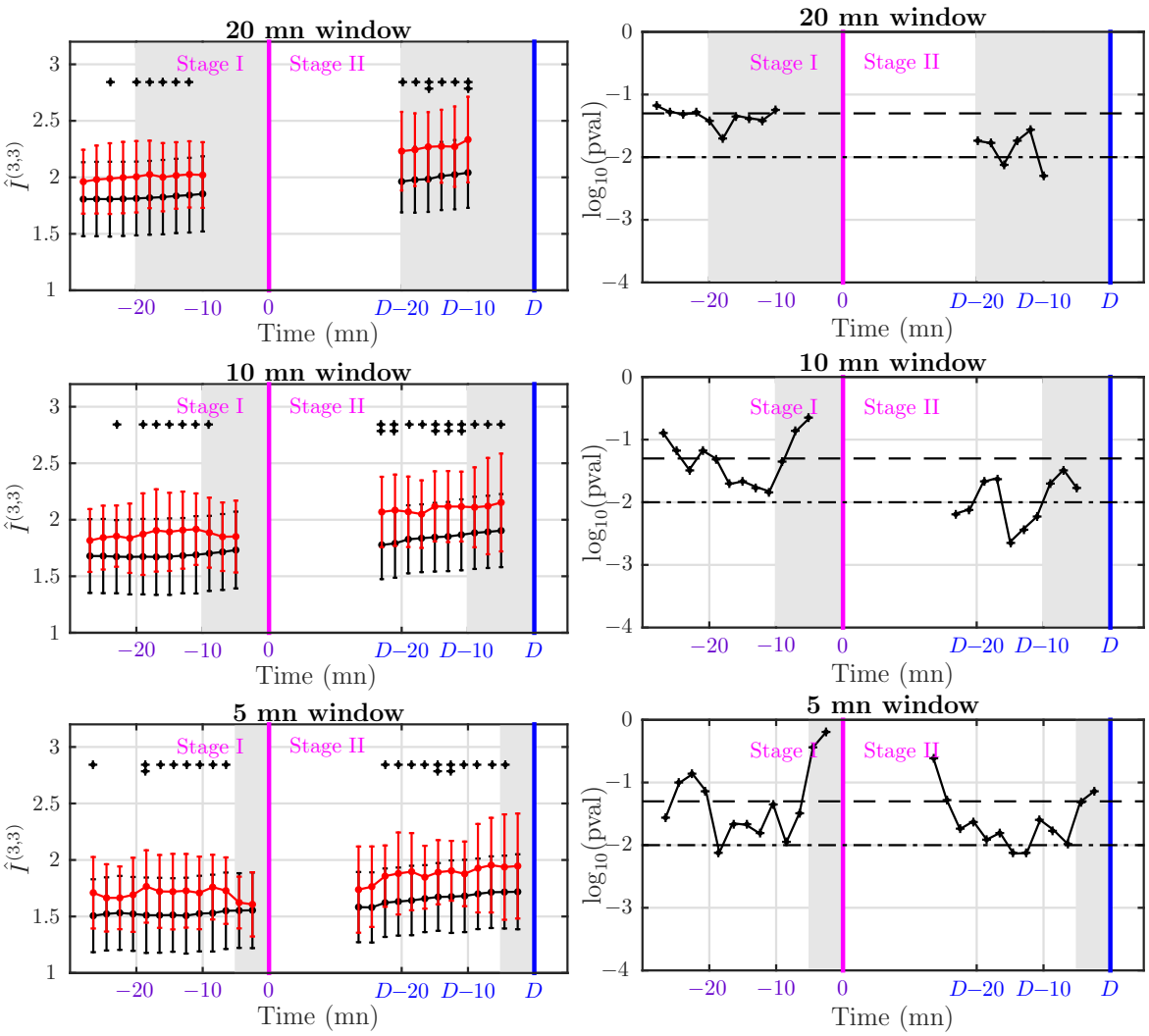

Figure 5: Left: average AMI $\hat{I}^{(3,3)}$ for normal (black) and abnormal (red) subjects in dataset II (delivery occuring more then 15mn after pushing started). AMI is computed in windows of size $T = 20$mn (first line), 10mn (second line) and 5mn (third line) shifted by 2mn. Vertical magenta line indicate the beginning of stage 2 (pushing) and vertical blue line indicate delivery. Right: corresponding p-value. Black + symbols in the AMI plot indicate a p-value lower than 5% (+) or lower than 1% (++). The gray-shaded region represents the time window used to compute the last value of AMI.

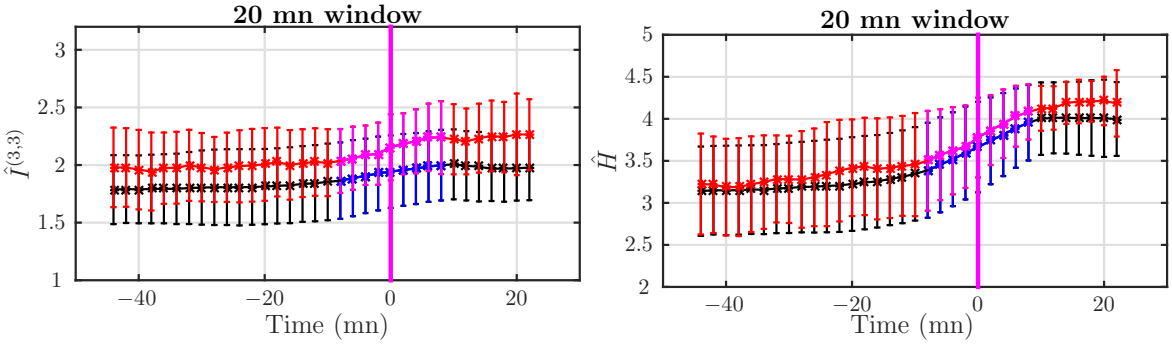

Figure 6: Average behavior of AMI $\hat{I}^{(3,3)}$ (left) and Shannon entropy $\hat{H}$ (right) for normal (black) and acidotic (red) subjects in dataset II (delivery occuring more than 15mn after pushing started). Quantities are computed in windows of size $T = 20$mn. Vertical magenta line indicate the beginning of stage 2 (pushing). We have used a different color code for windows spanning both stages I and II: blue for normal subjects and magenta for acidotic ones.

small scale than normal ones. We can relate this increase of the dependence structure of acidotic FHR to the short term variability and to its coupling with particular large scale patterns. For example, a sinusoidal FHR pattern [2], especially if its duration is long, should give a larger value of the AMI, because its large scale dynamics is highly predictable. As another example, we expect *variable decelerations* (with an asymmetrical V-shape) and *late decelerations* (with a symmetrical U-shape and/or reduced variability) to impact AMI differently. Of course, the choice of the embedding parameter τ is then crucial, and this is currently under investigation.

AMI and entropy rates depend on the dynamics as they operate on time-embedded vectors. AMI focuses on nonlinear temporal dynamics, while being insensitive to the dominant static information. AMI is independent of the standard deviation which on the contrary contributes strongly to the Shannon entropy. This explains why AMI performs better than entropy rates estimates, such as $\text{ApEn}^{(m)}$, $\text{SampEn}^{(m)}$ and $\hat{h}^{(m)}$, which depend also on the standard deviation.

5.2 Acidosis detection in second stage

The results reported for stage 2 show a severe decrease in performance of the five estimated quantities. Analysing stage 2 is far more challenging than analysing stage 1, which suggests that temporal dynamics in stage 2 differ notably for those of stage 1 [38] or simply that our database does not contain enough acidotic subject in that case. $\hat{I}^{(m,p)}$ achieves the best performance in terms of p-value and AUC; this clearly underlines that the analysis of nonlinear temporal dynamics is critical for fetal acidosis detection in stage 2. As in stage 1, the AMI is always larger in stage 2 for acidotic subjects than for normal subjects.

Although the Shannon entropy computed from the last 20mn of stage 2 before delivery does not show a clear tendency in Fig. 1 for dataset II, looking at Figure 6 clearly shows that $\hat{H}$ increases as labor progresses: this is probably related to the average increase of the number of decelerations, which is expected in both the normal and acidotic population.

$\text{SampEn}^{(2)}$ is also able to perform discrimination in stage 2. From these observations, one may envision the definition of a new estimator that would measure the auto mutual information using the Rényi order-2 entropy by applying eq.(12). Nevertheless, it should be emphasized that Rényi order-q entropy is lacking chain rule of conditional probabilities as soon as $q \neq 1$, which may jeopardize any practical use of such an estimator.

5.3 Probing the dynamics

Increasing the total embedding dimensions in AMI improves the performance in the detection of acidotic subjects, in both first and second stages. The best performance is found for different total embedding dimension in the two datasets. This suggests that FHR dynamics is different in each stage.

As seen in eq. (11), the Shannon entropy rate can be split in two contributions: one that depends only on static properties (the Shannon entropy, estimated by $\hat{H}$) and one that involves the signal dynamics (the auto mutual information, estimated by $\hat{I}^{(m,1)}$). By following the time evolution of these two parts, we were

able to relate Shannon entropy $\hat{H}$ with the evolution of the labor and AMI not only with the evolution of the labor, but also with possible acidosis. Looking at subjects for which the pushing phase is longer than 15mn, it clearly appears that all fetuses are affected by the pushing, as evidenced by a large increase of the Shannon entropy $\hat{H}$, and a small increase of AMI. Additionally, the increase of AMI is steeper for abnormal subjects than for normal ones, which may indicate different reactions to the pushing and can be related to specific pathological FHR patterns. When the pushing stage is long (dataset II), fetuses reported as acidotic are not showing any sign of acidosis until prolongated pushing. These fetuses appear as normal until delivery is near.

When acidosis develops during the first stage of labor, in dataset I, we observe clearly that while AMI increases steadily till delivery for healthy fetuses, it increases faster for acidotic ones. This suggests that acidotic fetuses in dataset I react to early labor, as early as one hour before pushing starts. This could not only indicate that some fetuses are prone to acidosis, but also may pave the way for an early detection of acidosis in this case.

Acknowledgments This work was supported by the LABEX iMUST (ANR-10-LABX-0064) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

References

[1] Chandraharan, E.; Arulkumaran, S. Prevention of birth asphyxia: responding appropriately to cardiotocograph (CTG) traces. *Best Pract. Res. Clin. Obstet. Gynaecol.* **2007**, *21*, 609–624.

[2] Ayres-de Campos, D.; Spong, C.Y.; Chandraharan, E.; Intrapartum Fetal Monitoring Expert Consensus Panel, F.I.G.O. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. *Int J Gynaecol Obstet* **2015**, *131*, 13–24.

[3] FIGO Subcommittee on Standards in Perinatal Medicine Guidelines for the Use of Fetal Monitoring. *Int. J. Gynaecol. Obstet.* **1987**, *25*, 159–167.

[4] Hruban, L.; Spilka, J.; Chudek, V.; Jank, P.; Huptych, M.; Bura, M.; Hudec, A.; Kacerovsk, M.; Kouck, M.; Prochzka, M.; Koreko, V.; Segetá, J.; imetka, O.; Mchurov, A.; Lhotsk, L. Agreement on intrapartum cardiotocogram recordings between expert obstetricians. *Journal of Evaluation in Clinical Practice* **2015**, *21*, 694–702.

[5] Alfirevic, Z.; Devane, D.; Gyte, G.M. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst Rev* **2006**, *3*, CD006066.

[6] Spilka, J.; Chudáček, V.; Koucký, M.; Lhotská, L.; Huptych, M.; Janků, P.; Georgoulas, G.; Stylios, C. Using nonlinear features for fetal heart rate classification. *Biomedical Signal Processing and Control* **2012**, *7*, 350–357.

[7] Haritopoulos, M.; Illanes, A.; Nandi, A., Survey on Cardiotocography Feature Extraction Algorithms for Fetal Welfare Assessment. In *XIV Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON), 2016, Paphos, Cyprus*; Kyriacou, E.; Christofides, S.; Pattichis, C.S., Eds.; Springer International Publishing: Cham, 2016; pp. 1187–1192.

[8] Akselrod, S.; Gordon, D.; Ubel, F.A.; Shannon, D.C.; Berger, A.C.; Cohen, R.J. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science* **1981**, *213*, 220–222.

[9] Gonçalves, H.; Rocha, A.P.; de Campos, D.A.; Bernardes, J. Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery. *Med Biol Eng Comput* **2006**, *44*, 847–855.

[10] Van Laar, J.; Porath, M.; Peters, C.; Oei, S. Spectral analysis of fetal heart rate variability for fetal surveillance: Review of the literature. *Acta Obstet. Gynecol. Scand.* **2008**, *87*, 300–306.

[11] Siira, S.M.; Ojala, T.H.; Vahlberg, T.J.; Rosén, K.G.; Ekholm, E.M. Do spectral bands of fetal heart rate variability associate with concomitant fetal scalp pH? *Early Hum Dev* **2013**, *89*, 739–742.

[12] Francis, D.P.; Willson, K.; Georgiadou, P.; Wensel, R.; Davies, L.; Coats, A.; Piepoli, M. Physiological basis of fractal complexity properties of heart rate variability in man. *J Physiol* **2002**, *542*, 619–629.

[13] Doret, M.; Helgason, H.; Abry, P.; Gonçalvés, P.; Gharib, C.; Gaucherand, P. Multifractal analysis of fetal heart rate variability in fetuses with and without severe acidosis during labor. *American Journal of Perinatology* **2011**, *28*, 259–266.

[14] Doret, M.; Spilka, J.; Chudáček, V.; Gonçalves, P.; Abry, P. Fractal Analysis and Hurst Parameter for intrapartum fetal heart rate variability analysis: A versatile alternative to Frequency bands and LF/HF ratio. *PLoS ONE* **2015**, *10*, e0136661.

[15] Echeverria, J.C.; Hayes-Gill, B.R.; Crowe, J.A.; Woolfson, M.S.; Croaker, G.D.H. Detrended fluctuation analysis: a suitable method for studying fetal heart rate variability? *Physiol Meas* **2004**, *25*, 763–774.

[16] Chudek, V.; Anden, J.; Mallat, S.; Abry, P.; Doret, M. Scattering Transform for Intrapartum Fetal Heart Rate Variability Fractal Analysis: A Case-Control Study. *IEEE Transactions on Biomedical Engineering* **2014**, *61*, 1100–1108.

[17] Magenes, G.; Signorini, M.G.; Arduini, D. Classification of cardiotocographic records by neural networks. *Proc. IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN)* **2000**, *3*, 637–641.

[18] Pincus, S.M.; Viscarello, R.R. Approximate entropy: a regularity measure for fetal heart rate analysis. *Obstet Gynecol* **1992**, *79*, 249–255.

[19] Pincus, S. Approximate entropy (ApEn) as a complexity measure. *Chaos* **1995**, *5 (1)*, 110–117.

[20] Lake, D.E.; Richman, J.S.; Griffin, M.P.; Moorman, J.R. Sample entropy analysis of neonatal heart rate variability. *Am J Physiol Regul Integr Comp Physiol* **2002**, *283*, R789–R797.

[21] Georgieva, A.; Payne, S.J.; Moulden, M.; Redman, C.W.G. Artificial neural networks applied to fetal monitoring in labour. *Neural Computing and Applications* **2013**, *22*, 85–93.

[22] Czabanski, R.; Jezewski, J.; Matonia, A.; Jezewski, M. Computerized analysis of fetal heart rate signals as the predictor of neonatal acidemia. *Expert Systems with Applications* **2012**, *39*, 11846–11860.

[23] Warrick, P.; Hamilton, E.; Precup, D.; Kearney, R. Classification of normal and hypoxic fetuses from systems modeling of intrapartum cardiotocography. *IEEE Transactions on Biomedical Engineering* **2010**, *57*, 771–779.

[24] Spilka, J.; Frecon, J.; Leonarduzzi, R.; Pustelnik, N.; Abry, P.; Doret, M. Sparse Support Vector Machine for Intrapartum Fetal Heart Rate Classification. *IEEE Journal of Biomedical and Health Informatics* **2016**, *PP*, 1–1.

[25] Dawes, G.; Moulden, M.; Sheil, O.; Redman, C. Approximate entropy, a statistic of regularity, applied to fetal heart rate data before and during labor. *Obstet Gynecol* **1992**, *80(5)*, 763–8.

[26] Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* **2000**, *278*, H2039–H2049.

[27] Grünwald, P.; Vitányistard, P. Kolmogorov complexity and information theory. With an interpretation in terms of questions and answers. *Journal of Logic, Language and Information* **2003**, *12*, 497–529.

[28] Lake, D.E. Renyi entropy measures of heart rate Gaussianity. *IEEE Transactions on Bio-Medical Engineering* **2006**, *53*, 21–27.

[29] Grassberger, P.; Procaccia, I. Estimation of the Kolmogorov-entropy From A Chaotic Signal. *Physical Review A* **1983**, *28*, 2591–2593.

[30] Kozachenko, L.F.; Leonenko, N.N. Sample Estimate of the Entropy of a Random Vector. *Problems of Information Transmission* **1987**, *23*, 95–101.

[31] Porta, A.; Bari, V.; Bassani, T.; Marchi, A.; Tassin, S.; Canesi, M.; Barbic, F.; Furlan, R. Entropy-based complexity of the cardiovascular control in Parkinson disease: Comparison between binning and k-nearest-neighbor approaches. *Conf Proc IEEE Eng Med Biol Soc (EMBC)* **2013**, pp. 5045–5048.

[32] Spilka, J.; Roux, S.; Garnier, N.; Abry, P.; Goncalves, P.; Doret, M. Nearest-neighbor based wavelet entropy rate measures for intrapartum fetal heart rate variability. *Conf. Proc. of the IEEE Eng. Med. Biol. Soc. (EMBC)* **2014**, pp. 2813–2816.

[33] Xiong, W.; Faes, L.; Ivanov, P.C. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Phys. Rev. E* **2017**, *95*, 062114.

[34] Costa, A.; Ayres-de Campos, D.; Costa, F.; Santos, C.; Bernardes, J. Prediction of neonatal acidemia by computer analysis of fetal heart rate and ST event signals. *Am J Obstet Gynecol* **2009**, *201*, 464.e1–464.e6.

[35] Georgieva, A.; Papageorghiou, A.T.; Payne, S.J.; Moulden, M.; Redman, C.W.G. Phase-rectified signal averaging for intrapartum electronic fetal heart rate monitoring is related to acidaemia at birth. *BJOG* **2014**, *121*, 889–894.

[36] Spilka, J.; Abry, P.; Goncalves, P.; Doret, M. Impacts of first and second labour stages on Hurst parameter based intrapartum fetal heart rate analysis. Computing in Cardiology Conference (CinC), 2014, pp. 777–780.

[37] Lim, J.; Kwon, J.Y.; Song, J.; Choi, H.; Shin, J.C.; Park, I.Y. Quantitative comparison of entropy analysis of fetal heart rate variability related to the different stages of labor. *Early Hum Dev* **2014**, *90*, 81–85.

[38] Spilka, J.; Leonarduzzi, R.; Chudáček, V.; Abry, P.; Doret, M. Fetal Heart Rate Classification: First vs Second Stage of Labor. Proc. Int. Workshop Biosignal Interpretation (BSI); , 2016.

[39] Granero-Belinchon, C.; Roux, S.; Garnier, N.; Abry, P.; Doret, M. Mutual information for intrapartum fetal heart rate analysis. Con Proc IEEE Eng Med Biol Soc (EMBC), 2017.

[40] Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters* **2002**, *89*, 068102.

[41] Doret, M.; Massoud, M.; Constans, A.; Gaucherand, P. Use of peripartum ST analysis of fetal electrocardiogram without blood sampling: a large prospective cohort study. *Eur J Obstet Gynecol Reprod Biol* **2011**, *156*, 35–40.

[42] Takens, F. Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence* **1981**, *4*, 366–381.

[43] Shannon, C. A Mathematical Theory of Communication. *The Bell System Technical Journal* **1948**, *27*, 388–427.

[44] Rényi, A. On measures of information and entropy. Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability, 1961, pp. 547–561.

[45] Kullback, S.; Leibler, R. On information and sufficiency. *Annals of Mathematical Statistics* **1951**, *22*, 79–86.

[46] Gomez, C.; Lizier, J.; Schaum, M.; Wollstadt, P.; Grutzner, C.; Uhlhaas, P.; Freitag, C.; Schlitt, S.; Bolte, S.; Hornero, R.; Wibral, M. Reduced predictable information in brain signals in autism spectrum disorder. *Frontiers in neuroinformatics* **2014**, *8*, 9.

[47] Faes, L.; Porta, A.; Nollo, G. Information Decomposition in Bivariate Systems: Theory and Application to Cardiorespiratory Dynamics. *Entropy* **2015**, *17*, 277–303.

[48] Granero-Belinchon, C.; Roux, S.G.; Garnier, N.B. Scaling of information in turbulence. *Epl* **2016**, *115*, 58003.

[49] Eckmann, J.P.; Ruelle, D. Ergodic-theory of Chaos and Strange Attractors. *Reviews of Modern Physics* **1985**, *57*, 617–656.

[50] Pincus, S.M. Approximate Entropy As A Measure of System-complexity. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2297–2301.

[51] Krstacic, G.; Krstacic, A.; Martinis, M.; Vargovic, E.; Knezevic, A.; Smalcelj, A.; Jembrek-Gostovic, M.; Gamberger, D.; Smuc, T. Non-linear analysis of heart rate variability in patients with coronary heart disease. *Computers in Cardiology* **2002**, *29*, 673–675.

[52] Richman, J.; Moorman, R. Time series analysis using approximate entropy and sample entropy. *Biophysical Journal* **2000**, *78*, 218A–218A.

[53] Lake, D. Improved entropy rate estimation in physiological data. *Conf. Proc. of the IEEE Eng Med Biol Soc (EMBC)* **2011**, pp. 1463–1466.

[54] Jaynes, E.T. Information Theory and Statistical Mechanics. Brandeis Lectures in theoretical Physics. Brandeis Summer Institute, 1962, Vol. 3.

[55] Singh, H.; Misra, N.; Hnizdo, V.; Fedorowicz, A.; Demchuk, E. Nearest neighbor estimates of entropy. *American Journal of Mathematical and Management Sciences* **2003**, *23*, 301–321.

[56] Kraskov, A.; Stogbauer, H.; Grassberger, P. Estimating mutual information. *Physical Review E* **2004**, *69*, 066138.

[57] Gao, W.; Oh, S.; Viswanath, P. Demystifying Fixed k-Nearest Neighbor Information Estimators. *CoRR* **2016**, *abs/1604.03006*.

[58] Freeman, R. Problems with intrapartum fetal heart rate monitoring interpretation and patient management. *Obstet Gynecol* **2002**, *100*, 813–26.

[59] Signorini, M.; Magenes, G.; Cerutti, S.; Arduini, D. Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings. *IEEE Transactions on Biomedical Engineering* **2003**, *50*, 365–374.

[60] Gonçalves, H.; Henriques-Coelho, T.; Bernardes, J.; Rocha, A.; Nogueira, A.; Leite-Moreira, A. Linear and nonlinear heart-rate analysis in a rat model of acute anoxia. *Physiol Meas* **2008**, *29(9)*, 1133–43.