Adaptive Multiscale Complexity Analysis of Fetal Heart Rate

H. Helgason*, P. Abry, P. Gonçalvès, Cl. Gharib, P. Gaucherand, and M. Doret

***Abstract*—*Per partum* fetal asphyxia is a major cause of neonatal morbidity and mortality. Fetal heart rate monitoring plays an important role in early detection of acidosis, an indicator for asphyxia. This problem is addressed in this paper by introducing a novel complexity analysis of fetal heart rate data, based on producing a collection of piecewise linear approximations of varying dimensions from which a measure of complexity is extracted. This procedure specifically accounts for the highly nonstationary context of labor by being adaptive and multiscale. Using a reference dataset, made of real *per partum* fetal heart rate data, collected *in situ* and carefully constituted by obstetricians, the behavior of the proposed approach is analyzed and illustrated. Its performance is evaluated in terms of the rate of correct acidosis detection versus the rate of false detection, as well as how early the detection is made. Computational cost is also discussed. The results are shown to be extremely promising and further potential uses of the tool are discussed. MATLAB routines implementing the procedure will be made available at the time of publication.**

***Index Terms*—Fetal heart rate (FHR) monitoring, multiscale approximations, network flow algorithms, per partum acidosis detection, time series complexity.**

I. Introduction

A. Per Partum Asphyxia Detection

PER PARTUM fetal heart rate (FHR) monitoring aims at reducing neonatal morbidity and mortality due to asphyxia. Intrapartum foetal asphyxia is responsible for 30% of cerebral palsy in term neonates [1]. Commonly, in making their decisions, obstetricians analyze FHR visually and follow the the International Federation of Gynecology and Obstetrics (FIGO)-guidelines [2] or the American College of Obstetricians and Gynecologists (ACOG) classification and guidelines [3]. This approach has shown to exhibit high sensitivity (high True Positive rate); however, at the price of very low specificity (high False Positive rate) to detect fetal asphyxia and subsequent cerebral palsy [4]. This results in a significant number of unnecessary operative deliveries (instrumental delivery, caesarean) performed for fetuses showing at birth, hence *a posteriori*, no sign of asphyxia (False Positives) [5]. Compared to spontaneous vaginal delivery, operative deliveries are associated with higher short- and long-term morbidity risk, both for the newborn and for the mother [6]–[8]. For that reason, providing obstetricians with a robust and efficient statistical index aiming at assisting them in detecting *per partum* asphyxia efficiently constitutes an important goal and challenge in daily clinical practice.

B. FHR Analysis

FHR is mainly regulated by the autonomic nervous system whose activity highly depends on blood oxygenation content and pressure. This makes quantitative and systematic analysis of *per partum* FHR meaningful for predicting asphyxia. The automated and statistical analysis of FHR during labor, however, appears challenging, especially since FHR during labor is highly nonstationary and continuously evolving until delivery. Moreover, FHR is, in many cases, characterized by *decelerations*, occurring more or less regularly, sometimes induced by contractions, and consisting of sharp drops in heart rate that differ radically from the normal evolution of the FHR variability (which classically refers to high-frequency fluctuations of heart rate). Both the FHR variability and the sharp onset of decelerations may, however, have energies in the same frequency bands. Moreover, in forming their decisions, obstetricians take into account both the number and the shapes of the decelerations and the short-time variability of the FHR, hence making the use of local information over of a wide range of timescales of the data. This prevents one from using classical tools in heart rate variability analysis that strongly rely on data stationarity and/or mostly concentrate uniformly along time on the short-time fluctuations, as is the case for methods based on standard spectral analysis (see, e.g., [9]). Instead, techniques that account for multiple timescales have recently been applied to fetal and adult heart rate, such as fractal and multifractal analysis [10]–[13], and local regularity analysis [14], [15]. Other appealing approaches aim at measuring the complexity of the data using entropy (see, e.g., [16], [17]) or nonlinear data modeling (see, e.g., [18]), and are, therefore, grounded on concepts very different from the ones underlying the method being presented here. Time-varying modeling of adult heart rate has been proposed in [19], but is rather dedicated to smooth changes that may turn inappropriate for analysis of FHR where abrupt and transient changes occur.

Manuscript received November 29, 2010; revised January 29, 2011; accepted February 15, 2011. Date of publication March 3, 2011; date of current version July 20, 2011. The work of P. Abry was supported by the *Young Research Team* award granted by Foundation del Duca, Académie des Sciences, Institut de France, 2007. The work of H. Helgason was supported by the Physics Department, ENS de Lyon, Lyon, France. *Asterisk indicates corresponding author.*

*H. Helgason is with the School of Electrical Engineering, Kungliga Tekniska högskolan (KTH), Stockholm SE-100 44, Sweden (e-mail: hannes.helgason@ee.kth.se).

P. Abry is with the Department of Physics, Centre National de la Recherche Scientifique (CNRS), École Normale Supérieure de Lyon, Lyon 69364, France (e-mail: patrice.abry@ens-lyon.fr).

P. Gonçalvès is with the Department of Computer Sciences, Institut National de Recherche en Informatique et en Automatique (INRIA), École Normale Supérieure de Lyon, Lyon 69364, France (e-mail: paulo.goncalves@ens-lyon.fr).

C. Gharib, P. Gaucherand, and M. Doret are with the Hôpital Femme-Mère-Enfant, Hospices Civils de Lyon, Université Lyon, Lyon 69364, France (e-mail: gharib@pop.univ-lyon1.fr; pascal.gaucherand@chu-lyon.fr; muriel.doret@chu-lyon.fr).

Digital Object Identifier 10.1109/TBME.2011.2121906

Another important challenge is that the methodology has to be extendible for online processing of FHR in order to actually help obstetricians and not only provide off-line postanalysis. Consequently, the computational cost and memory requirements have to allow for real-time analysis to be implemented in delivery rooms.

C. Goals and Contributions

In this context, this contribution aims at proposing a novel approach for measuring the complexity of FHR data, extracted from the computation of a family of (near-continuous piecewise linear) approximations of FHR data. The originality of the proposed method stems from the fact that it renews classical *trend versus variability* heart rate analysis in the following respect: 1) it is multiscale (allows for nonuniform time-resolution) and adaptive (data driven), two key properties that efficiently address the nonstationary and widely varying nature of *per partum* FHR; and 2) it *jointly* analyzes trend and variability. Moreover, the method is originally grounded on a network flow algorithm enabling fast computations of the proposed *per partum* FHR approximations. The principles and algorithms related to the measure of complexity are detailed in Section III. The method is then applied to a collection of real *per partum* FHR time series, carefully constituted by obstetricians and described in Section II. The potential of the tool in decreasing the False Positive rate (hence, the number of unnecessary operative deliveries) and in performing early detections of *per partum* fetal acidosis is discussed in Section IV.

II. Dataset

A. Data Recording

This case-control study was conducted using a dataset collected in the years 2000–2007 at the Department of Obstetrics of the French Academic Hospital *Hôpital Femme-Mère-Enfant* (Bron, France). FHR analysis based on fetal ECG monitoring is routinely performed for fetus with high risk of *per partum* asphyxia, as indicated by suspicious FHR observed on initial cardiotocographic recordings. The FHR is recorded during labor with either a STAN S21 or a STAN S31 monitor (STAN, Neoventa Medical, Moelndal, Sweden), using a scalp electrode with 12-bit resolution and 500-Hz sampling rate. The recordings typically have durations ranging from 30 min to several hours.

B. Dataset

Forty-seven FHR recordings, among patients with no obstetrical pathology, were selected by the obstetricians conducting this study, according to two independent information: umbilical cord acid-base status and FHR pattern classifications. Umbilical cord acid–base status is evaluated *a posteriori* (i.e., after birth) and yields a classification of the fetuses into: 1) *healthy;* or 2) acidotic, to which we will refer, hereafter, as *unhealthy*. FHR pattern classification is based on the FIGO criteria [2]. It is conducted by the attending obstetrician in charge during the delivery process and leads to classifying the FHR pattern as *normal* or *abnormal*, the latter being a key element leading to the decision to perform an operative delivery. We split the available database into three groups by combining the *a posteriori* information related to the fetus health status to the *a priori* FHR pattern classification (also see Table I):

TABLE I
DATA DESCRIPTION

Group	Status	# subj.	Dur. (min)	Gestational Weeks	Umb. Cord pH
FIGO-TN	Healthy	15	225 ± 114	$39^{6/7}$ ± 10 (days)	7.34 ± 0.03 $[7.30, 7.40]$
FIGO-FP	Healthy	17	206 ± 125	$39^{6/7}$ ± 10 (days)	7.32 ± 0.02 $[7.30, 7.40]$
FIGO-TP	Unhealt.	15	181 ± 146	40 ± 8 (days)	7.00 ± 0.03 $[6.95, 7.05]$

1) *FIGO-TN*: 15 healthy fetuses, normal FHR pattern;
2) *FIGO-FP*: 17 healthy fetuses, abnormal FHR pattern;
3) *FIGO-TP*: 15 unhealthy fetuses, abnormal FHR pattern.

The third group is labeled *FIGO-TP*, for True Positive, to indicate acidotic (unhealthy) status of the fetus and the classification of the FHR pattern as abnormal. The same reasoning holds for *FIGO-FP*, for False Positive, and *FIGO-TN*, for True Negative. These three groups can be gathered into the two classes *healthy* (FIGO-TN and FIGO-FP, 32 subjects) and *unhealthy* (FIGO-TP, 15 subjects). The grouping provides us with a performance benchmark: using FIGO classification, obstetricians achieved on this dataset a 100% True Positive rate at the price of a $17/(17+15) \approx 53\%$ False Positive rate. Keeping in mind that False Positive detection can lead to operative delivery that should be avoided, the goal of the present contribution is to evaluate to which extent the classification based on our proposed methodology can decrease the False Positive rate while preserving a 100% True Positive rate.

C. Preprocessing

The STAN system records and stores lists of time occurrences of R-peaks $\{\tau_k, k = 1, \ldots, K\}$ (in seconds). A common practice in heart rate analysis (see, e.g., [9], [20]) is to transform this sequence of times into a regularly sampled heart rate time series $(y_n)_{n=1}^N$ (in beats per minute, B/M) constructed by interpolating the set of points $\{(\tau_k, 60(\tau_{k+1} - \tau_k)^{-1}), k = 1, \ldots, K-1\}$ and resampling regularly at times $(t_n)_{n=1}^N$. Because the frequency content of the data is essentially concentrated in the range $[0, 2.5]$ Hz, the resampling frequency is set to $F_s = 8$ Hz (approximately twice the Shannon sampling rate). The interpolation scheme also accounts for missing data. It has been carefully checked that varying either F_s or the interpolation scheme does not affect the results reported later.

D. Sliding-Window Analysis

In daily clinical practice, obstetricians implicitly perform a time-sliding analysis of the data, considering the characteristics of FHR within the few previous minutes and tracking their evolution along time. Once a start of acidosis is suspected and decision has been made, an operative delivery should be performed within the next 15–30 min. At their suggestions, we perform a sliding time window analysis of the data. To comply with the

time from decision to operative delivery, time window lengths are chosen close to $15/2 = 7.5$ min; the dyadic specificity of our algorithm led us to consider a 8.5-min window length, corresponding to the closest sample size equal to a power of two. This could be easily changed. For the sake of simplicity, results are presented for nonoverlapping windows, but in practice one could overlap them to get closer to continuous monitoring. We analyzed the last 90 min of the FHR recordings before birth (some of the recordings were shorter than this).

III. Graph-Based Adaptive Multiscale Approximations and Measure of Complexity

A. Sequence of Multiscale Approximations of FHR

1) Approximations: The adaptive multiscale (near-) continuous piecewise approximations of FHR time series $(y_n)_{n=1}^N$ will be of the form

$$\widetilde{y}_W(t) = \sum_{v \in W} g_v(t) \tag{1}$$

where W denotes a finite collection of indices labeling *templates* g_v assumed to have disjoint time supports and to be of the form

$$g_v(t) = (a(t - t_{I,0}) + b) \cdot 1_I(t), \quad v = (a, b, I) \tag{2}$$

with a, b, and $I := [t_{I,0}, t_{I,1})$ referred to as the *slope*, *offset*, and *support* of g_v. Furthermore, the set $(g_v)_{v \in W}$ is required to satisfy *continuity constraints*, as described in Section III-B.

2) Discretization: For the purpose of constructing fast algorithms, we will consider *dyadic interval* template supports, which when working on the time interval $[0, 1)$ (the data can always be rescaled to satisfy this) are of the form $I_{j,k} = [k2^{-j}, (k+1)2^{-j})$, where $j \geq 0$ is the *scale index* and $k \in \{0, \ldots, 2^j - 1\}$ is the *dyadic index* for the interval. Slopes and offsets of templates at scale 2^{-j} are discretized as $a = l\Delta b\, 2^j$, $b = m\Delta b$, $l, m \in \mathbb{Z}$, for some predetermined offset step size $\Delta b > 0$; see Fig. 1(left) for a schematic diagram of two templates. Consequently, all template endpoints necessarily fall on a discrete grid $\{(k2^{-J_1}, m\Delta b) : k = 0, \ldots, 2^{J_1}, m = M_0, M_0 + 1, \ldots, M_1\}$, where 2^{-J_1} is the finest scale and M_0 and M_1 are determined by the range of offsets to be considered. Note that here the range of available templates can be further restricted according to *a priori* available physiological information, such as, natural bounds on heart rate, maximum rate of change in heart rate, etc. A more detailed discussion about discretization and approximation theoretical issues can be found in [21].

3) Sequence of Best Constrained Fits: We define the "best" *L-constrained approximation* (or *fit*) $\widehat{y}_L$ as the approximation that minimizes the mean square difference between the data among all admissible approximations $\widetilde{y}_W$ with $|W| = L$

$$C(L) := \min_{|W|=L} \|y - \widetilde{y}_W\|_{\ell_2}^2 = \|y - \widehat{y}_L\|_{\ell_2}^2. \tag{3}$$

In this setting, $L = |W|$, the number of templates in (1), acts as a measure of sparsity, with a natural interpretation as the approximation's *dimension*. For small L, $\widehat{y}_L$ mainly models the global trends in the data ($L = 1$ amounts to approximating data with a single linear trend). When increasing L, $\widehat{y}_L$ will typically first refine the approximation of the largest decelerations or variations, and then start progressively capturing the small-scale FHR variability. In this sense, $C(L)$ can be regarded as an L-dependent (multiscale) measure of the complexity of the data. Illustrations of a collection of approximations $\widehat{y}_L$ for various L are shown in Fig. 2 for an FHR tracing from the dataset described in Section II. Examples of $(L, C(L))$ curves, together with best constrained approximations are illustrated in Fig. 3. A description of an algorithm for efficiently obtaining $\widehat{y}_L$ and $C(L)$ is postponed to Section III-B.

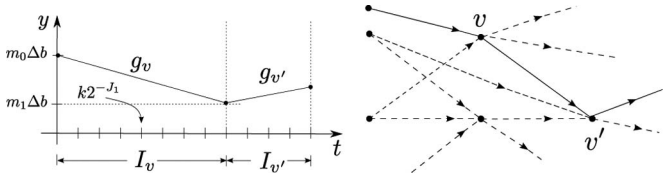

Fig. 1. *Left diagram:* two templates satisfying continuity constraint. g_v has time support I_v, offset $m_0\Delta b$, and slope $a = (m_1 - m_0)\Delta b\,|I_v|^{-1}$; all dyadic intervals have endpoints of the form $k\,2^{-J_1}$ where the finest time scale is 2^{-J_1}. *Right diagram:* depiction of a template graph; vertices represent templates, broken lines show connectivities, solid lines show a path.

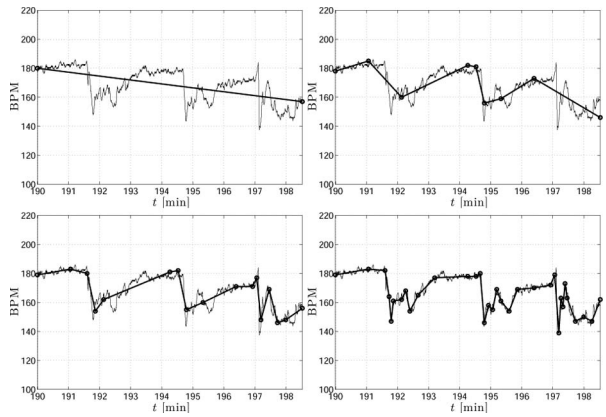

Fig. 2. Examples of FHR tracings (for a FIGO-FP) and best constrained approximations $\widehat{y}_L$ for four different choices of $L = |W|$: $L = 1, 8, 16, 32$.

Rather than considering (3) for a sequence of Ls, a classical strategy could be to seek for an "optimal" L according to some criteria. One classical approach that arises in model selection (see, e.g., [22]) would be to consider complexity functionals of the form $\mathcal{L}_\lambda(\widetilde{y}_W \mid y) = \|y - \widetilde{y}_W\|_{\ell_2}^2 + \lambda|W|$, for fixed $\lambda > 0$, and without any constraint on $|W|$. Since, in general, $C(L)$ is a monotonously decreasing function of L, it is easy to show that this complexity functional is actually minimized at $\widehat{y}_{L^*}$ where $L^* := arg\min_L (C(L) + \lambda L)$. However, choosing a parameter λ for $\mathcal{L}_\lambda$ (or picking another relevant penalization functional) seems difficult due to the large variety of possible shapes FHR tracings one can have. Instead, it is proposed here to use the full curve $(L, C(L))$ since it avoids both the difficulties of choosing a model selection scheme, and it also contains more information than restricting to a single value of L. Perhaps more importantly, as supported by the FHR analysis in [12] and [13], there is evidence of scaling in FHR tracings that suggests that one should, in fact, look at how $C(L)$ changes with varying L.

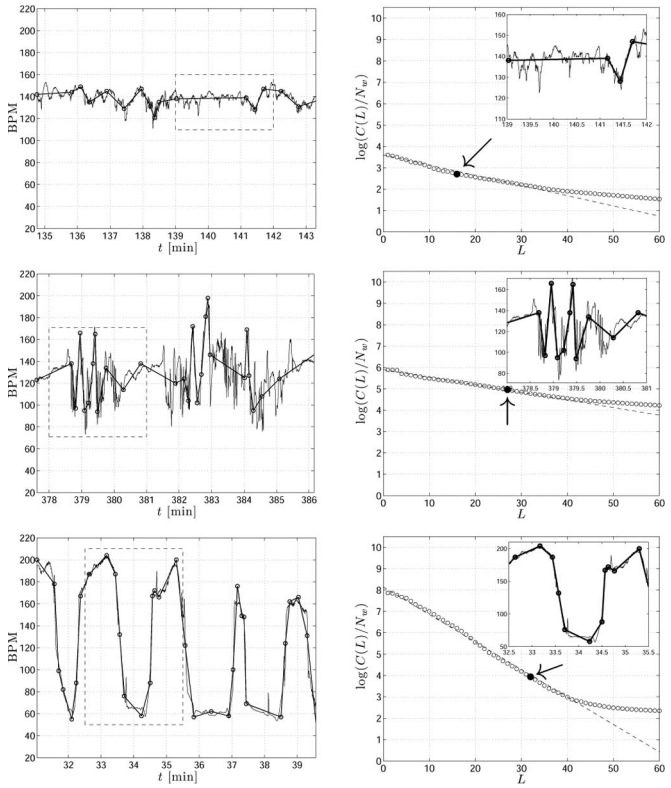

Fig. 3. Some FHR tracings (left) and corresponding plots of $(L, \log(C(L)/N_w))$ (right) with least-squares fits $\kappa L + \theta$ in the range $L = 1, \ldots, 32$ (dashed lines). N_w is the number of samples in the 8.5-min time window being analyzed (see Section IV-A). Inserts show close-ups of the overlaid constrained best fits (arrows point at the corresponding Ls). The method's adaptive capability is illustrated on arbitrarily chosen examples of FIGO-TN (top), FIGO-FP (middle), FIGO-TP (bottom).

In the present contribution, we claim that the shape of the curve $(L, C(L))$ provides information that can help to characterize FHR and can be used in acidosis detection in fetus during labor.

B. Constrained Shortest Path Algorithm

This section presents an algorithm for computing $C(L)$ and finding the corresponding "best" L-approximation $\widehat{y}_L$ in (3) for a range of different and fixed dimensions L. The algorithm elaborates on ideas presented in [21] and [23] and is based on a graph structure and a constrained shortest path algorithm akin to the one introduced in [24].

Although the constrained shortest path problem is, in general, NP-complete, the special topology of our graph—namely, that it is *directed* and *acyclic*—allows it to be solved in polynomial time, i.e., with a number of operations proportional to the number of edges in the graph; this latter depending on the discretization of templates and continuity constraints. Because the algorithm proposed here elaborates on the *constrained shortest path algorithm* framework, theoretically devised and analyzed in [21], [23], and [24], it benefits of the properties shown to hold therein, which fully justifies that it provides *best* approximations [as defined in (3)] as well as the evaluation of its computational complexity. Key elements are overviewed as follows.

Algorithm 1 Calculation of Local Fits

J_0 and J_1 denote the coarsest and finest scales
Initialization Step I (done over the finest scale index J_1)
for $k = 0, \ldots, 2^{J_1} - 1$
    Calculate the terms in (8) and store them
Initialization Step II (recursion over coarser scales)
for $j = J_1 - 1, \ldots, J_0$
    for $k = 0, \ldots, 2^j - 1$
        Calculate and store the terms in (8)
Calculate Fits: (loop over templates)
for each template $g_v(t)$ of the form (2):
    Use (6), (7) and the precomputed values in (8)
    to calculate (4), the local fit of template v

1) Template Costs: We define the *template cost* (or the local fit) $c_v(y)$ for template g_v, given data y, by

$$c_v(y) := \|y - g_v\|_I^2 = \|y\|_I^2 - 2\langle y, g_v\rangle_I + \|g_v\|_I^2 \tag{4}$$

where $\langle f, g\rangle_I := \sum_{t_n \in I} f(t_n)g(t_n)$ and $\|f\|_I^2 := \langle f, f\rangle_I$, and writing I instead of I_v (for cleaner notation). Using (1) together with the constraint that the templates in the collection $(g_v)_{v\in W}$ have disjoint time supports yields

$$\|y - \widetilde{y}_W\|_{\ell_2}^2 = \sum_{v\in W} \|y - g_v\|_{I_v}^2 = \sum_{v\in W} c_v(y) \tag{5}$$

so that the squared difference between the data and an approximation $\widetilde{y}_W$ is simply the sum of the template costs $(c_v(y))_{v\in W}$. The terms in (4) can be broken down as follows:

$$\begin{aligned}\|g_v\|_I^2 &= a^2\langle t^2, 1\rangle_I + (2ab - 2a^2 t_{I,0})\langle t, 1\rangle_I \\ &\quad + (a^2 t_{I,0}^2 - 2abt_{I,0} + b^2)|I| \end{aligned}\tag{6}$$

$$\langle y, g_v\rangle_I = a\langle y, t\rangle_I + (b - at_{I,0})\langle y, 1\rangle_I \tag{7}$$

where $|I|$ is the number of samples in $I = [t_{I,0}, t_{I,1})$. The terms

$$\langle y, 1\rangle_I \quad \langle y, t\rangle_I \quad \|y\|_I^2 \quad |I| \quad \langle t, 1\rangle_I \quad \langle t^2, 1\rangle_I \tag{8}$$

can, therefore, be calculated once for every dyadic interval and then (6) and (7) used to calculate local fits (4) for different slopes and offsets. The property $\langle u, v\rangle_{I''} = \langle u, v\rangle_I + \langle u, v\rangle_{I'}$, for $I'' = I \cup I', I \cap I' = \emptyset$, can be used to calculate the terms (8) recursively, going from coarse to fine scales. The full calculation procedure is sketched in Algorithm 1.

2) Template Graph: Having a discretized set of templates, we consider approximations (1) where the collection $(g_v)_{v\in W}$ satisfies certain *a priori* chosen constraints: 1) the support of $\widetilde{y}$ is the whole interval $[0, 1]$; 2) the templates $(g_v)_{v\in W}$ have disjoint time supports; and 3) for every two adjacent templates, g_v and $g_{v'}$ where $t = t'$ is their time of junction, $|g_v(t') - g_{v'}(t')|$ is small (i.e., the approximation is near continuous).

The whole set of discretized templates g_v can be regarded as vertices v in a *template graph*, where the directed edges (v, v'), connecting vertices v and v', are determined by the constraints earlier; Fig. 1(right) is a schematic diagram showing part of a template graph. Let the cost associated with vertex v be defined as $c_v(y)$ (i.e, the template cost for g_v). A path in the graph is a collection W of connected vertices originating

from a template starting at time $t = 0$ and terminating at a template ending at time $t = 1$. Note that the set of paths in the graph exactly corresponds to the set of approximations (1) we wish to consider. Moreover, the squared difference (5) for an approximation $\widetilde{y}_W$ is the *cost* of the path W in the graph, defined by the sum of the costs of the vertices it visits. The *length* L of a path is the number of vertices (templates) it visits. Thus, solving the minimization problem (3) for a fixed L corresponds to finding the path of length L in the template graph that has the minimum cost, that is, it amounts to solving a constrained shortest path problem.

The graph is directed and vertices on any path in the graph are visited exactly once, i.e., the graph is acyclic and, hence, contains no loops. This fact is of extreme importance since it gives access to fast network flow algorithms for solving (3).

3) Constrained Best Paths: Because the graph is directed and acyclic, the vertices (v_i), $i = 1, \ldots, |V|$ can be ordered using a *topological ordering* [25] such that for each edge (v_i, v_j), $i < j$ is satisfied. This can be done easily by moving along the time axis from left to right, taking time steps corresponding to the smallest template scale under consideration. At each step, the templates are labeled starting from the current time position; these templates can be ordered freely since they do not connect to each other due to connectivity constraint 2). Since templates starting at later times are assigned a larger label, this gives us a topological ordering. Vertex v_i is referred to directly by its topological order i. Let $d_i(\ell)$ be the tentative cost of the shortest path up to vertex i constituted by exactly ℓ edges (i.e., paths visiting exactly ℓ vertices). The lengths ℓ range from 0 to $L_{\max}$, where $L_{\max}$ is the maximum number of edges allowed in a path and is either determined by the smallest template scale considered or chosen *a priori*. Define a source node v_0, who connects to all templates (vertices) starting at time $t = 0$ (all paths W start from the source node). Denote by $\text{pred}_i(\ell)$ the vertex that precedes vertex i in the tentative best path of length ℓ from the source node to vertex i. The algorithm used to extract the constrained best path of length L is given in Algorithm 2. Its output $d_i(\ell)$ consists of the lowest path cost up to vertex i constituting ℓ edges and $\text{pred}_i(\ell)$ is the vertex that precedes vertex i on this path. Thus, Algorithm 2 finds the constrained best cost for all vertices in the graph in just one sweep over the set of vertices. Note that this algorithm is a simple extension of the classical (unconstrained) shortest path algorithm and is justified in an almost identical fashion (see [25] for detailed discussion about shortest path algorithms).

4) Computational Costs: For Algorithm 1, Step I requires at most $6N$ additions and $3N$ multiplications, where N denotes the number of samples y_n. In Step II, a recursion over coarser scales is performed where each step in the inner loop requires a constant number of additions of order $O(2^{J_1} - 2^{J_0})$. Therefore, the computational complexity for Algorithm 1 is $O(N)$. Algorithm 2 is of *fixed* and *known* computational complexity that depends only on the size and topology (discretization and constraints) of the template graph. Its overall computational and memory cost scale as $O(L_{\max}|E|)$ and $O(|V| \times L_{\max})$, respectively, where $|V|$ and $|E|$ denote the number of vertices and edges in the graph. Thus, solving the optimization problem (3) to calculate $C(L)$ and the corresponding best approximations $\widehat{y}_L$ for $L = 1, \ldots, L_{\max}$ has low computational cost and memory requirements.

Algorithm 2 Calculation of Best Constrained Fits

Initialize distance labels:
Set $d_0(\cdot) = 0$ and $d_i(\cdot) = \infty$ for $i = 1, \ldots, |V|$
Update distance labels in topological order:
for $i = 1, \ldots, |V|$
  Let A_i be the set of edges going out from vertex v_i.
  for all $(v_i, v_j) \in A_i$
    for $l = 1, \ldots, L_{\max}$
      if $d_j(l) > d_i(l-1) + c_{v_j}$
        $d_j(l) = d_i(l-1) + c_{v_j}$
        $\text{pred}_j(l) = i$.

IV. Acidosis Detection Using FHR Data

The procedure for computing $C(L)$ (see Section III-B) is now applied to the dataset described in Section II with the aim of evaluating the benefits for *per partum* fetal acidosis detection.

A. Exploiting $C(L)$

The shape of the curve $(L, \log(C(L)/N_w))$ (where N_w is the number of samples in the analyzed time window), as shown in Fig. 3, is viewed as a functional representation of the complexity of the analyzed data and appears interesting for discriminating unhealthy from healthy subjects. Given the limited number of patients available in this study, blindly feeding all the values of points on the curve into a standard classification technique is unreasonable. Instead, we chose to model the curve parametrically. Inspection of the curves for all patients in the dataset suggests to use piecewise linear models: two-piece (four parameters) and one-piece (two parameters) models were investigated. Interestingly, and surprisingly, it has been observed *a posteriori* that the best classification performance are obtained when curves are characterized by the slope and offset parameters in a least-squares linear fit $\kappa L + \theta$ to the curve in the range $L = 1, \ldots, 32$. Notably, the use of the change point in values of L in a 2-piece model did not improve classification. Further exploration of these issues will be conducted on larger datasets (see Section V).

We retain the model $\kappa L + \theta$ in the remaining analysis, also because its parameters receive simple physiological interpretations. The intercept θ corresponds to the sample variance of the data, and thus measures the variability of the FHR, a criterion widely used by obstetricians; very low variability indicates bad health for the fetuses. The global variability alone is usually not a sufficient index and obstetricians further refine it by considering long-term ($\sim$60 s) versus short-term ($\sim$3.75 s) variabilities [26]. This is notably the case in situations such as *per partum* FHR analyzed here, where various phenomena occurring during the delivery process (contractions notably) complicate the use of well-established time scale ranges. In the modeling proposed here, as explained in Section III-A, $C(L)$ consists of the remaining variability after an approximation using L segments

has been removed from the data. Hence, it measures a parameter dependent variability of the data. The tool proposed here offers two refinements to the usual long-term versus short-term variability analysis: 1) it is multiscale since the time segments in the approximations can have different lengths; and 2) it is adaptive, since the set of analyzing scales are not chosen *a priori*, as is the case in, e.g., spectral analysis—instead, relevant scales are selected by the data as a reaction to constraining L, the number of segments. In that second respect, the slope parameter κ receive a natural interpretation: the smaller the absolute value of κ (note $\kappa \leq 0$), the heavier the short term variability. Steep slopes κ give indication that there are large scale trends in the FHR, such as decelerations. These interpretations are illustrated in Fig. 3.

B. Online FHR Monitoring

For each subject, the curve $(L, \log(C(L)/N_w))$ is computed independently for each of the M available time windows, giving a set of characterization parameters $\{(\kappa_m, \theta_m), m = 1, \ldots, M\}$ (M depends on the length of the FHR recording). Fig. 4 (top plot) shows examples of time evolutions of (κ_m, θ_m) for the three groups. Fig. 4 (middle plot) shows a scatter plot of the parameters for all subjects and all windows. These plots indicate that large absolute values of κ_m and θ_m provide evidence of unhealthy subjects (i.e., the FIGO-TP group) and also that the κ_m and θ_m of the FIGO-FP group differ from those of the FIGO-TN group. Also, it can be noticed that the two parameters appear more correlated for the unhealthy subjects.

Each data recording begins at an arbitrary time decided by the obstetrician and ends at the time of delivery. To abide with practical clinical conditions, the following analysis will consider quantitative detection rules that monitor the parameter values (κ_m, θ_m) sequentially; i.e., as m grows, sliding from one window to the next. The detection rule will either decide at some window $m = m^*$ that the subject is unhealthy, or continue until the end of the FHR recording, in which case the scheme would classify the subject as healthy.

C. Acidosis Detection

1) One-Parameter Detection Rules: Fig. 4 (bottom plot) shows a scatter plot of $\bar{\kappa}$ and $\bar{\theta}$ that denote the extreme absolute values of κ_m and θ_m, respectively, taken over the whole FHR recording for each subject. The plot indicates that unhealthy subjects form a cluster away from the healthy subjects (i.e., the FIGO-TN and FIGO-FP groups). Kruskal–Wallis ranksum tests applied separately to $\bar{\kappa}$ and $\bar{\theta}$, to compare the healthy and unhealthy classes, yield, in both cases, p-values of the order 10^{-5}, hence validating their potential for distinguishing healthy from unhealthy subjects. Moreover, the dashed lines in Fig. 4 (bottom plot) show the *ideal* thresholds C_κ and C_θ chosen so that the True Positive rate (correct detection) is 100%, when $\bar{\kappa}_M > C_\kappa$ OR $\bar{\theta}_M > C_\theta$. Using these *ideal* thresholds, the corresponding False Positive rate is around 65% using only κ_m, and close to 80% using only θ_m. Both of these detection rules give worse performance than the available 53% FIGO benchmark. However, a detection rule based on $\bar{\kappa}_M > C_\kappa$ AND $\bar{\theta}_M > C_\theta$, yields False Positive rate around 47%, much closer to the benchmark. This detection rule corresponds to observing a point (κ_m, θ_m) in the bottom right region of Fig. 4 (bottom plot). This *a posteriori* analysis indicates that both κ and θ are meaningful for FHR analysis and should be considered jointly.

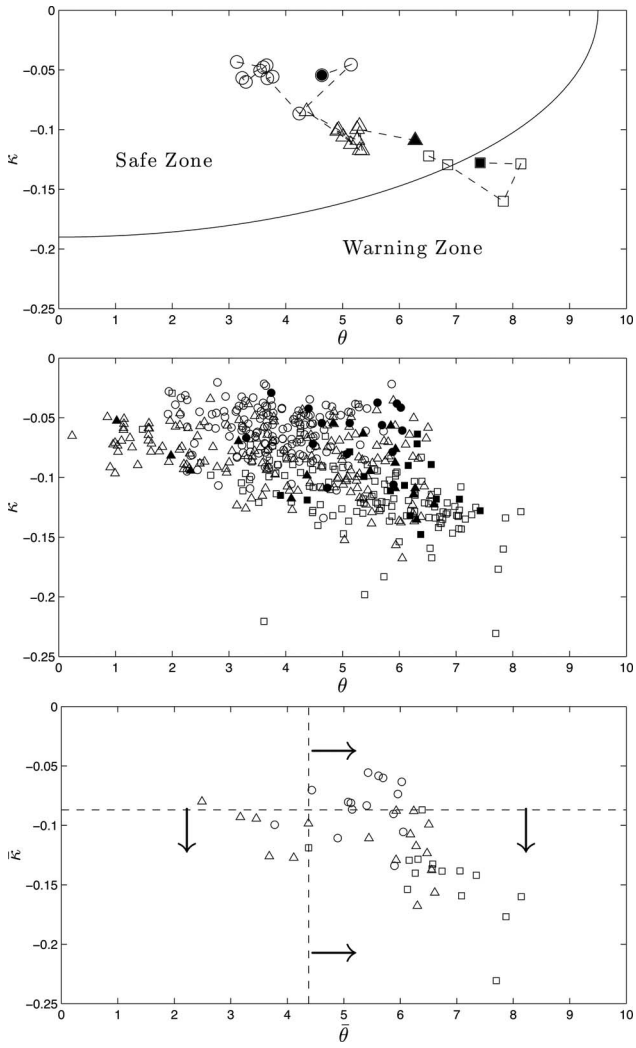

Fig. 4. Top: time evolution of (κ_m, θ_m) for three subjects; elliptic curve demonstrates a split of domain into warning and safe zones (see Section IV-C). Middle: scatter plot of (κ_m, θ_m) for all the subjects and all windows; solid symbols indicate last time window. Bottom: scatter plot of the extreme values $\bar{\kappa}$ and $\bar{\theta}$; broken lines depict the ideal thresholds giving a 100% True Positive rate using one-parameter detection rules for the dataset, arrows point in the direction of detection regions. (FIGO-TN: circle "◯;" FIGO-FP: triangle "△" FIGO-TP: box "□").

2) Two-Parameter Detection Rules: Fig. 4 (top and middle) indicates that the parameter values (κ_m, θ_m) "live" in different regions of the (κ, θ) plane for the two classes of healthy and unhealthy subjects: healthy subjects concentrate in the upper-left part, referred to as the *Safe Zone*; unhealthy subjects are spread in the lower-right part, the *Warning Zone*. Searching for classification boundaries is the subject of a large corpus in the statistical learning literature (see, e.g., [22] for an overview), but because this case study relies on a small dataset, the recourse to such tools is difficult here (note also that each subject contributes to

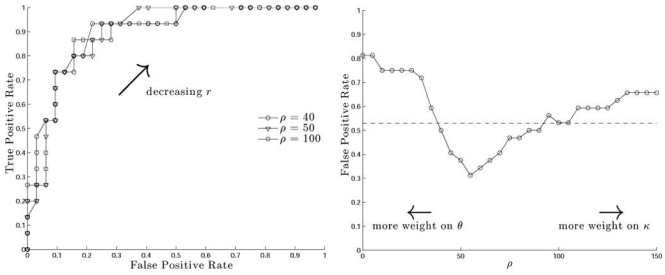

Fig. 5. *Left:* ROC curves for different values of ρ, the shape parameter for the elliptical boundary $\theta^2 + (\rho\kappa)^2 = r^2$. *Right:* False Positive rate as a function of ρ in the elliptical boundary $\theta^2 + (\rho\kappa)^2 = r^2$; r is chosen *a posteriori* to give 100% True Positive rate; horizontal broken line corresponds to the FIGO-FP rate.

TABLE II
PROPORTION OF ALARM TIMES FOR IDEAL BOUNDARIES (ROUNDED IN %)

$M - m^*$	0	1	2	3	4	5	6	7	8	9	10
$\rho = 40$	13	13	27	20	7	0	7	0	0	0	13
$\rho = 50$	7	7	27	27	7	7	7	0	0	0	13
$\rho = 100$	7	7	27	20	0	13	7	0	13	0	7

measurements in a series of time windows, putting us in a situation of multiple comparisons). Instead, we will use heuristic elliptical boundaries in the (κ, θ) plane, given by the equation $\theta^2 + (\rho\kappa)^2 = r^2$, to explore classification performance. The parameter ρ is the ratio of minor and major axes of the ellipse and, hence, controls the shape of the boundary. Small ρ gives more weight to the intercept parameter θ, while conversely, large ρ puts emphasis on the slope parameter κ. The parameter r controls the distance of the boundary from origin, hence keeping ρ fixed, r controls the size of the *Safe Zone*. Fig. 4 illustrates an example of the boundary for $(\rho, r) = (50, 9.5)$.

Receiver operating characteristic (ROC) curves, obtained by varying r, are plotted in Fig. 5 for different values of ρ. They show that comparable performance can be obtained for various choices of the pair (ρ, r) and, moreover, a satisfactory closeness to the upper left corner: for each ρ, a value of r can be found, such that the False Positive rate does not exceed 53% while retaining a 100% True Positive rate.

To go further, we now select, for each ρ, the value of r enabling the lowest False Positive rate, while insuring a True Positive rate of 100%, in accordance with the obstetrician request that not a single unhealthy subject is missed. The corresponding False Positive rate is represented in Fig. 5(right plot) as a function of ρ. For large ρ, emphasis is put on κ and the False Positive rate tends to 65%; conversely, for small ρ, emphasis is on θ and the False Positive rate tends to 80%; in agreement with the results reported for the one parameter detection rules. The important observation is that for a wide range of ρ, the False Positive rate goes down to 30%, hence improving considerably the performance for fetal acidosis detection. This constitutes the first major result of this study.

3) Alarm Time: We now investigate the (conservative) *alarm time,* which is the position m^* of the time window, within which the parameters (κ_m, θ_m) are furthest away from the elliptical boundary; that is, we let m^* be where $\theta_m^2 + (\rho\kappa_m)^2$ is maximum. This implies that the decision is made *at latest* $(M - m^*) \times 8.5$ min before delivery. Table II reports, as a function of $M - m^*$ for different ρ, the proportion of unhealthy subjects for which the maximum of $\theta_m^2 + (\rho\kappa_m)^2$ occurs in window m^* (the table shows rounded numbers, e.g., writing 7 instead of $1/15 \approx 6.7$). For only around 10% of the subjects, detection actually takes place in the last time window. For more than 50% of the subjects, detections occurs for $M - m^* = 2$ or 3, i.e., from 17 to 25 min prior to the operative delivery. Interestingly, in 13% of the cases, detection could have been performed as early as $M - m^* = 8$ or 10, i.e., roughly 70 to 90 min before delivery, and actually as soon as the monitoring was started, whereas doctors normally expect that FHR complexity should change only very gradually to become clearly apparent only in the last 30 min before birth.

Several possible interpretations of the results can be proposed. First, acidosis setup could be more progressive than suggested by clinical studies, based on indirect signs as fetal pH cannot be continuously monitored. Second, rather than identifying fetal acidosis, this method could actually detect fetuses challenged by hypoxia, the step systematically preceding fetal asphyxia and acidosis. This would be very interesting for clinicians who aim at reducing fetal exposure to asphyxia. Third, some fetuses may undergo alternate phases of acidosis and recovery, seen as a precursor of a permanent acidosis phase. The formulation of such hypotheses is regarded by the obstetricians as motivating outcome of this complexity analysis of FHR data and will be further investigated. Therefore, this potential early detection capability constitutes the second major result of this study.

V. CONCLUSION AND PERSPECTIVES

We have proposed and evaluated a novel and flexible methodology for the analysis of FHR during labor. It is based on analyzing complexity of FHR data from a collection of multiscale and adaptive near-continuous piecewise linear approximations, obtained using a fast network flow algorithm.

The use of the proposed measure of complexity for detecting fetal *per partum* acidosis has been evaluated using a dataset of real FHR data that were collected *in situ* and selected by obstetricians as representative for healthy and unhealthy fetuses. This method could potentially outperform the current FIGO benchmark. The possibility of early acidosis detection is also of particular interest on the medical side and demands further investigation.

The analysis tools developed here can be applied to any FHR monitoring techniques delivering beat-to-beat RR interval information. In clinical situations, this currently excludes techniques other than scalp electrode ECG, such as abdominal ECG or cardiotocograph that only provide medium term ($\sim$30 beats) averaged R wave to R wave interval information. However, recent technological progresses in such techniques may enable to deliver beat-to-beat data. This is opening further potential uses of the analysis tools developed in this paper.

To further comfort the potential of this new approach of FHR analysis, the complexity characterization procedure will be applied to a much larger dataset, currently under constitution, consisting of about 4 000 subjects. In that context, it will be

coupled with more elaborated classification procedures based on modern statistical learning techniques. The procedure can also be further tuned to FHR data by incorporating more specific physiological and medical knowledge into the design of the templates and the graph underlying the best-constrained approximation search. We also plan to use this scheme to help locating and characterizing the *per partum* FHR decelerations and, hence, to specifically focus the analysis onto them. At the moment they are accounted for in the complexity measurements globally and among other FHR data features.

At the methodological level, further investigations aiming at providing understanding of the nature of the time series complexity measured here might be obtained from the analysis of reference processes, such as fractional Brownian motion. Another research direction is to refine the framework by considering different templates or using other measures of local fits. It seems also necessary to tailor the methodology better to online monitoring, where one could imagine leveraging the sequential property of the network flow algorithm being used; it marches forward in time and does not need to look at data from the past that makes the algorithm particularly well suited for online processing of data. Another exciting direction is to extend this framework for direct processing of RR intervals; this would not require the data to be interpolated and resampled. This could be fruitfully compared to the noteworthy method proposed in [19], which relies on ARMA-based history dependent point process modeling of the RR interval series.

MATLAB toolbox implementing the methodology will be made publicly available at the time of publication. The parameter settings for the template graph used in the analysis in Section IV will also be made available.

ACKNOWLEDGMENT

The authors would like to thank Neoventa for technical help in RR interval extraction for this data.

REFERENCES

[1] D. S. Reddihough and K. J. Collins, "The epidemiology and causes of cerebral palsy," *Aust. J. Physiother.*, vol. 49, pp. 7–12, 2003.

[2] FIGO, "Guidelines for the use of fetal monitoring," *Int. J. Gynaecol. Obstet.*, vol. 25, pp. 159–167, 1987.

[3] G. A. Macones, G. D. Hankins, C. Y. Spong, J. Hauth, and T. Moore, "The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: Update on definitions, interpretation, and research guidelines," *Obstet. Gynecol.*, vol. 112, no. 3, pp. 661–666, 2008.

[4] K. B. Nelson, J. M. Dambrosia, T. Y. Ting, and J. K. Grether, "Uncertain value of electronic fetal monitoring in predicting cerebral palsy," *N. Engl. J. Med.*, vol. 334, no. 10, pp. 613–619, 1996.

[5] Z. Alfirevic, D. Devane, and G. M. Gyte, "Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour," *Cochrane Database Syst. Rev.*, vol. 19, no. 3, p. CD006066, 2006.

[6] M. B. Landon, J. C. Hauth, K. J. Leveno, C. Y. Spong, S. Leindecker, M. W. Varner, A. H. Moawad, S. N. Caritis, M. Harper, R. J. Wapner, Y. Sorokin, M. Miodovnik, M. Carpenter, A. M. Peaceman, M. J. O'Sullivan, B. Sibai, O. Langer, J. M. Thorp, S. M. Ramin, B. M. Mercer, and S. G. Gabbe, "Maternal and perinatal outcomes associated with a trial of labor after prior cesarean delivery," *N. Engl. J. Med.*, vol. 351, no. 25, pp. 2581–2589, 2004.

[7] J. M. Alexander, K. J. Leveno, J. Hauth, M. B. Landon, E. Thom, C. Y. Spong, M. W. Varner, A. H. Moawad, S. N. Caritis, M. Harper, R. J. Wapner, Y. Sorokin, M. Miodovnik, M. J. O'Sullivan, B. M. Sibai, O. Langer, and S. G. Gabbe, "Fetal injury associated with cesarean delivery," *Obstet Gynecol.*, vol. 108, no. 4, pp. 885–890, 2006.

[8] R. Silver, M. B. Landon, D. J. Rouse, K. J. Leveno, C. Y. Spong, E. A. Thom, A. H. Moawad, S. N. Caritis, M. Harper, R. J. Wapner, Y. Sorokin, M. Miodovnik, M. Carpenter, A. M. Peaceman, M. J. O'Sullivan, B. Sibai, O. Langer, J. M. Thorp, S. M. Ramin, and B. M. Mercer, "Maternal morbidity associated with multiple repeat cesarean deliveries," *Obstet Gynecol.*, vol. 107, no. 6, pp. 1226–1232, 2006.

[9] J. V. Laar, M. Porath, C. Peters, and S. Oei, "Spectral analysis of the fetal heart rate variability for fetal surveillance: Review of the literature," *Acta Obstetrica et Gynecol.*, vol. 87, pp. 300–306, 2008.

[10] K. Kiyono, Z. R. Struzik, N. Aoyagi, and Y. Yamamoto, "Multiscale probability density function analysis: Non-Gaussian and scale-invariant fluctuations of healthy human heart rate," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 1, pp. 95–102, Jan. 2006.

[11] P. C. Ivanov, "Scale-invariant aspects of cardiac dynamics," *IEEE Eng. Med. Biol. Mag.*, vol. 26, no. 6, pp. 33–37, Nov./Dec. 2007.

[12] P. Abry, H. Helgason, P. Gonçalvès, E. Pereira, P. Gaucherand, and M. Doret, "Multifractal analysis of ECG for intrapartum diagnosis of fetal asphyxia," in *Proc. IEEE Int. Conf. Acoust. Speech Sig. Proc. (ICASSP)*, 2010, pp. 566–569.

[13] M. Doret, H. Helgason, P. Abry, P. Gonçalvès, C. Gharib, and P. Gaucherand, "Multifractal analysis of fetal heart rate variability in fetuses with and without severe acidosis during labor," *Amer. J. Perinatol.*, vol. 28, no. 4. pp. 259–266, Apr. 2011.

[14] Z. Struzik, "Revealing local variability properties of human heartbeat intervals with the local effective Hölder exponent," *Fractals*, vol. 9, no. 1, pp. 77–93, 2001.

[15] T. Nakamura, H. Horio, and Y. Chiba, "Local Hölder exponent analysis of heart rate variability in preterm infants," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 1, pp. 83–88, Jan. 2011.

[16] M. G. Signorini, G. Magenes, S. Cerutti, and D. Arduini, "Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 365–374, Mar. 2003.

[17] M. Ferrario, M. G. Signorini, G. Magenes, and S. Cerutti, "Comparison of entropy-based regularity estimators: Application to the fetal heart rate signal for the identification of fetal distress," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 1, pp. 119–125, Jan. 2006.

[18] W. Jarisch and J. S. Detwiler, "Statistical modeling of fetal heart rate variability," *IEEE Trans. Biomed. Eng.*, vol. 27, no. 10, pp. 582–589, Oct. 1980.

[19] R. Barbieri, E. C. Matten, A. A. Alabi, and B. E. M., "A point-process model of human heartbeat intervals: New definitions of heart rate and heart rate variability," *Amer. J. Physiol Heart Circ. Physiol.*, vol. 288, pp. H424–H435, 2005.

[20] P. Hopkins, N. Outram, N. Lofgren, E. Ifeachor, and K. Rosen, "A comparative study of fetal heart rate variability analysis techniques," in *Proc. 28th IEEE EMBS Conf.*, 2006, pp. 1784–1787.

[21] H. Helgason, "Nonparametric detection and estimation of highly oscillatory signals," Ph.D. dissertation, Dept. Applied Comput. Math., California Inst. Technol., Pasadena, CA, 2008.

[22] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York: Springer-Verlag, 2009.

[23] E. J. Candès, P. Charlton, and H. Helgason, "Detecting highly oscillatory signals by chirplet path pursuit," *Appl. Comput. Harmon. Anal.*, vol. 24, pp. 14–40, 2007.

[24] H. C. Joksch, "The shortest route problem with constraints," *J. Math. Anal. Appl.*, vol. 14, pp. 191–197, 1966.

[25] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows. Theory, Algorithms and Applications*. New York: Prentice-Hall, 1993.

[26] G. S. Dawes, C. W. G. Redman, and J. H. Smith, "Improvements in the registration and analysis of fetal heart rate records at the bedside," *J. Obstet. Gynaecol.*, vol. 92, pp. 317–325, 1985.

Authors' photographs and biographies not available at the time of publication.