Monitoring Cardiac Stress Using Features Extracted From S_1 Heart Sounds

Jonathan Herzig*, Amitai Bickel, Arie Eitan, and Nathan Intrator

Abstract—It is known that acoustic heart sounds carry significant information about the mechanical activity of the heart. In this paper, we present a novel type of cardiac monitoring based on heart sound analysis. Specifically, we study two morphological features and their associations with physiological changes from the baseline state. The framework is demonstrated on recordings during laparoscopic surgeries of 15 patients. Insufflation, which is performed during laparoscopic surgery, provides a controlled, externally induced cardiac stress, enabling an analysis of each patient with respect to their own baseline. We demonstrate that the proposed features change during cardiac stress, and the change is more significant for patients with cardiac problems. Furthermore, we show that other well-known ECG morphology features are less sensitive in this specific cardiac stress experiment.

Index Terms—Biomedical signal processing, cardiology, laparoscopy, morphological features, phonocardiography.

I. INTRODUCTION

PATIENTS undergoing noncardiac surgery are at increased risk of cardiovascular morbidity and mortality [1]. Every year, more than 33 million patients undergo noncardiac surgery in the United States alone. Approximately 4% of these patients experience perioperative complications [2]. Cardiovascular complications arising from noncardiac surgery exceeds 1 million patients worldwide [3], and these complications account for one third of perioperative deaths [4]. In the coming decades, it is predicted that perioperative complications will double due to the increasing size of the elderly population in most developed countries [2].

Monitoring cardiac stress is a standard procedure in critical care medicine. It is most commonly performed via electrocardiography (ECG). ECG monitoring alone is not adequate to detect ischemia in real time in the intensive care unit or in an intraoperative setting [5]–[7]. Specifically, conventional visual ECG monitoring for the detection of transient ST-segment changes is inaccurate [7].

Another way to monitor cardiac status is by analyzing the acoustic signals of the heart. Heart sounds (HS) are generated by mechanical cardiac activity—blood flow and the closure of valves. HS can change their timing, magnitude, and morphol-

ogy due to different cardiac events. These changes happen due to the complex interplay between pressure gradients in the atria, ventricles, and arteries. The resulting nonstationary signal indirectly reflects the physiological attributes of the subject [8]. The HS signal of a healthy heart is composed of two distinct components—the first HS, S_1, and the second HS, S_2. S_1 occurs at the end of the isometric contraction period, during systole, and S_2 occurs after the isovolumetric relaxation period, during diastole [9].

The pulmonary system contributes to the modulation of cardiovascular mechanical activity. Respiratory changes in pleural pressure, arterial resistance, and venous return can have significant effects. These effects lead to significant differences between the properties of S_1 and S_2, which occur during inspiration and expiration [10].

The field of automated analysis of HS signals requires work at several levels, from signal processing to pathology diagnosis. The most important avenues of research that are considered in this paper are listed below:

Noise removal: Removal of various types of noise (sound from lungs, background noise, and friction between the recording device and the skin), which can distort the features of basic HS components and decrease the quality of further analysis. Various methods have been used to separate HS signals from lung sounds, such as the manipulation of weight scales in wavelet multiresolution decomposition [11], [12] and singular spectrum analysis [13]. Other denoising methods include moment segmentation analysis [14] and multiple HS signal averaging [12], [15].

Segmentation: Segmentation of HSs into cardiac cycles and the partitioning of those cycles into HS components (S_1, systolic phase; S_2, diastolic phase). Methods for segmentation often involve choosing a signal representation, such as the homomorphic transform [16], [17], continuous wavelet transform [18], [19], Hilbert transform [20], or Stockwell transform [21], [22]. A threshold is set on the output signal to detect S_1 and S_2 events. The use of an external ECG signal can assist segmentation, as S_1 occurs after the QRS complex [15], [23].

Accurate localization of S_1: Some applications, e.g., tracking blood pressure changes and monitoring respiration, demand an accurate S_1 localization in addition to the basic segmentation of HS. This can be accomplished by using ensemble averaging [15] or Shannon entropy [13], [24].

This paper presents a general physiological monitoring system for cardiac stress, which is based on HSs. Features are extracted from the S_1 HSs during baseline (BL) activity and during stress, and these features are compared so that each patient serves as their own control. This type of monitoring can be performed while patients undergo surgery with anesthesia or

while participating in any other activity with a cardiac risk. The system is noninvasive and portable; therefore, it could also be used for home monitoring.

The framework presented in this paper is applied to HS signals recorded during laparoscopic surgeries. The immediate cardiovascular function changes that may occur due to CO_2 insufflation are characterized by a decrease in cardiac output, cardiac index, and stroke volume (decreased venous return), and an increase in systemic vascular resistance. This pattern is the result of the interaction between increased intra-abdominal pressure, neurohumoral responses, and the effects of absorbed CO_2 [25]. The effects of CO_2 (like mild respiratory acidosis) should be controlled by mechanical ventilation, and the proper evacuation of the gas at the end of surgery.

It should be noted that hemodynamic cardiovascular changes during laparoscopic operations depend on many factors, such as the patient's posture during abdominal insufflation, their hemodinamic status before insufflation, general medical health etc. Our study population included patients undergoing laparoscopic cholecystectomy, while the patient is in the antitrendelenburg (head up) position, expecting more adverse cardiovascular changes during creation of positive pressure pneumoperitoneum. Such inhomogeneity in subjects' response makes the chosen framework sufficiently challenging for a realistic study on cardiovascular changes under cardiac stress.

II. METHODS

The following computational framework was developed to identify the features of physiological changes that occur during and after cardiac stress. Monitoring these changes is critical for the prevention of cardiovascular complications during surgery. Early detection of these complications can help the surgeon to make an informed decision during surgery. In some cases, the procedure should be stopped to prevent further complications, and in other cases, longer hospital admission should be provided after the procedure is finished.

It is desirable to extract features for physiological changes that are comparable across patients. However, patients are heterogeneous. Each patient has a different cardiovascular characteristic behavior. Furthermore, some patients have no cardiac diseases, and others suffer from a variety of diseases with different physiological manifestations. We wish to apply a monitoring technique that takes into account the normal cardiac functioning of the patient and measures the deviation from normal function. In particular, we need the system to be useful for patients with a healthy heart as well as for those who suffer from a variety of cardiac problems. The novel feature of our method is that it addresses this problem by extracting features that compare the patients current cardiac state with the BL cardiac activity recorded previously. Our method, therefore, takes into account the normal cardiac activity of the individual.

We have analyzed recordings of HSs from several patients at BL activity and during cardiac stress, and we have extracted several candidate features that appear to distinguish the BL state from cardiac stress. Of these features, the two most significant features are presented here. Figs. 1 and 2 show recordings from
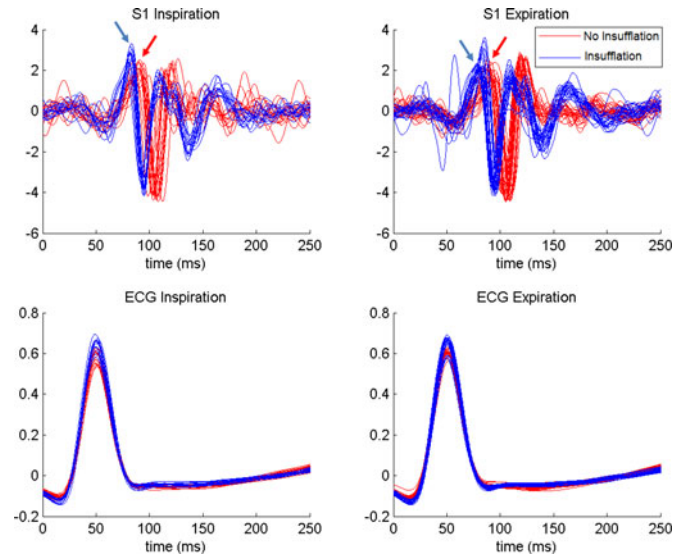

Fig. 1. Multiple S_1 HSs (upper row) and ECG recordings (lower row) for a single patient (patient ID: PA) that were recorded during laparoscopic surgery. The recording site for the HSs was the second left intercostal space (2L). HSs recorded before insufflation are shown in red. HSs recorded during insufflation are shown in blue. Inspiration samples are on the left, and expiration samples are on the right. ECG samples are all aligned such that R wave onset is at 50 ms. HSs onset portrays the heart electrical conductivity velocity. We see that the insufflation and noninsufflation phases have different average onsets, indicated respectively by the blue and red arrows. It is also notable that the ECG samples for the different phases are practically the same, showing little variance.
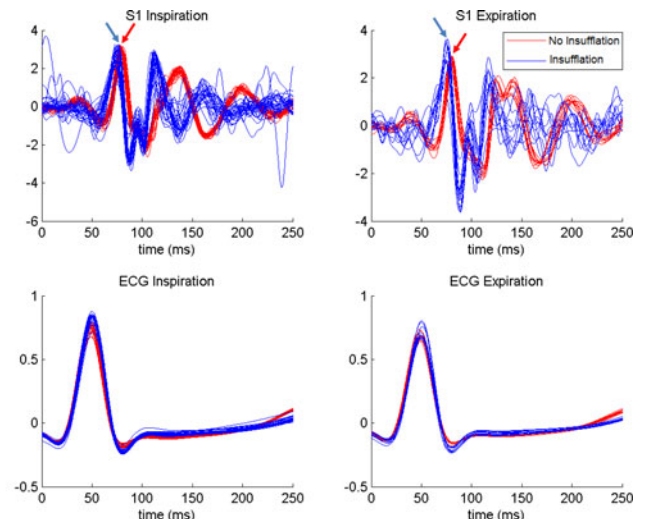

Fig. 2. Multiple HSs and ECG recordings with the same layout and recording site as Fig. 1 but for a different patient (patient ID: SUA). This example demonstrates the different morphologies that the insufflation and noninsufflation phases can have. Note that S_1 onset, indicated respectively by the blue and red arrows, is similar for the two phases, as is the ECG morphology.

patients that demonstrate the extensive deviation of HS features despite the similarity of the ECG readings. The first feature measures the difference in S_1 sound morphology between the BL and cardiac stress states. This nonparametric feature captures the physiological changes that are caused by cardiac mechanical activity, without addressing specific diagnostics. For
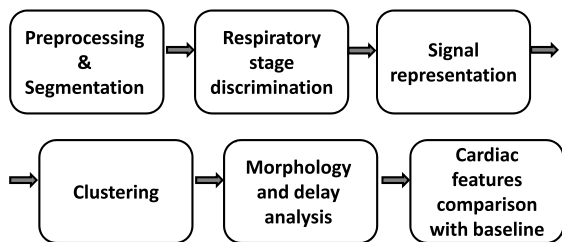
Fig. 3. Computational framework.

instance, morphology of recordings during insufflation presented in Fig. 2 may indicate split S_1 HSs [26] which results from asynchronous closure of mitral (M1) and tricuspid (T1) valves. However, this feature may be applicable to numerous cardiac problems that may arise during stress, such as cardiac ischemia and congestive heart failure. The second feature measures the cardiac electromechanical activation time. This represents the temporal delay in S_1 onset following the QRS wave (RS_1). This feature is linked with a specific physiological behavior, and it is known to have significantly different values for patients with impaired left ventricular contractility [27] and for patients who have experienced systolic heart failure [28]. Other features that were extracted and found to be less significant were measures of HS morphology deviations between the inspiration and expiration phases of respiration. These are related to the health of the heart and to intrathoracic pressure.

The details presented here refer to the setup for a noncardiac surgery, but the framework can be applied to any procedure with cardiac risk.

For each patient, two sets of S_1 HSs were recorded a BL set, which contains HSs recorded under anesthesia before the beginning of surgery, and a "monitoring state" set (MS). MS can contain HSs recorded during surgery, or HSs recorded after surgery, while the patient is still under anesthesia. Hierarchical clustering was applied to the BL set [23]. The most significant cluster centers were used as a template representing the physiological cardiac state under no cardiac stress. Clustering was also applied to all sets of HSs for outlier rejection. This is another novel feature of this paper; only the samples contained in the most significant clusters were kept for further analysis. For each set, the average delay relative to the BL template was calculated. The mean squared error (MSE) value for each set, relative to the BL template, was calculated as well to measure the change in signal morphology. Features representing the relationship between the two sets were then extracted from these calculations.

The computational framework consists of the building blocks described in Fig. 3 (described below in detail).

A. Preprocessing and Segmentation

HSs were recorded using a digital data acquisition system. The system is consisted of piezoelectric contact transducers, which were placed in the same locations for all patients. More details of the phonocardiogram recording are provided in a previous report [29]. The input to this stage is a vibro-acoustic signal consisting of multiple cardiac cycles that are to be monitored for cardiac stress. This signal can be recorded at any stage of the surgery. The output of this stage will be a set of S_1 sounds of the MS).

Except for producing the MS, the BL should be produced once during the monitoring procedure for each patient. This set consists of S_1 sounds recorded following the start of anesthesia. The motivation is to generate a template for patient-specific "normal" cardiac activity, which can be highly variable across patients. For this reason, each patient is assigned a specific BL. We assume that following the start of anesthesia, minimal cardiac stress is induced from the procedure itself, so we can refer to this stage as a BL, that is specific to the patient's cardiac functioning.

The acquired vibro-acoustic signal was digitally filtered by a Chebishev type-I IIR bandpass filter of order 4, with a passband between 20 and 75 Hz. The least noisy recording channel on this passband was manually selected for further analysis. The HS signal was then partitioned by using a simultaneously recorded ECG signal. The peaks of the ECG-QRS complexes were used as reference points. The S_1 sound was defined as the segment beginning 50 ms before the R-wave peak, and ending 250 ms later. These preprocessing parameters were selected as described previously [23]. This paper addresses monitoring by extracting the features of S_1 segments because most of the HS energy is concentrated in this component. S_1 segments are standardized after extraction. Segments with peaks significantly below or above average are considered to be invalid segments and are filtered out.

B. Inspiration and Expiration Discrimination

The respiratory cycle modulates the HS morphology: Heart beats that follow high breathing pressure during expiration are morphologically different from beats following low breathing pressure during inspiration [10]. We would like to find a compact BL representation for each patient, which best describes the cardiac activity for that patient. Therefore, we discriminate between S_1 sounds collected during expiration and inspiration, and we extract different compact BLs for each phase. The discrimination process is detailed below. Further analysis is applied separately for these two respiratory phases. The output of this stage includes BL and MS sets that contain only the segments of the chosen respiratory phase (inspiration or expiration).

To discriminate between the phases, we make use of the fact that the respiratory signal is generated by the movement of the chest during respiration. This signal can be extracted from the HS signal [10] by filtering with a bandpass filter in the range 0–0.5 Hz. The breathing signal is divided into respiratory cycles. The lowest point of the cycle corresponds to the end of expiration and the beginning of inspiration, and the highest point corresponds to the end of inspiration and the beginning of expiration. The respiratory cycle is mapped to the range of 0–360, the sinusoidal cycle. Inspiration corresponds to values of 0–180, and expiration to values of 180–360. Fig. 4 shows an example for a patient's breathing signal with the described mapping.
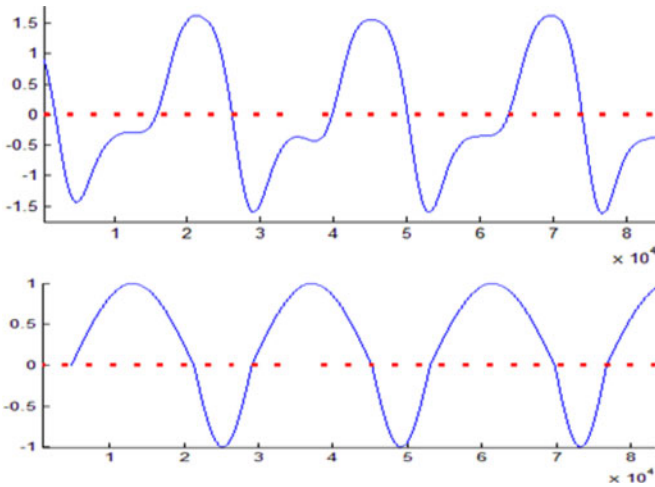
Fig. 4. (Top) three respiratory cycles. The lowest point of the cycle corresponds to the end of expiration and beginning of inspiration, and the highest point to the end of inspiration and the beginning of expiration. (Bottom) mapping of the respiratory cycle to the range 0–360. The graph shows a sinus of this map. The red intervals are times of S_1 beats.

C. Signal Representation

The choice of representation method for S_1 sound is fundamental for the analysis framework. Different representations highlight different features of the sound and may lead to different clustering results [30]. We examine three types of signal representations (which correspond to three types of estimators): Time domain, envelogram, and time–frequency representations.

The inputs to the signal representation stage are the MS and BL datasets. The outputs are these sets after transformation to a signal representation. The same transformation should be applied to both sets. We detail the transformation process for one set because it is similar for the other set: The input is a set of N S_1 sounds, $B = \{b_1, b_2, \ldots, b_N\}$, where b_i is the representation of an S_1 HS during a single cardiac cycle. All signals were recorded during the same phase of surgery. Furthermore, all signals were recorded during the same phase of the respiratory cycle (inspiration or expiration). The signals were given in a time-domain representation. Extracting the different signal representations from the detailed input was performed in the following way.

Time-domain representation: This is the raw representation of the data. Therefore, $B^{TD} = B$, where B^{TD} denotes the set of signals under the time-domain representation.

Hilbert envelogram representation: This representation carries less information, but it is also less noisy than the time-domain representation [20]. Every signal $b_i(t) \in B$ is transformed by Hilbert transform, to produce the envelogram, $b_i' = E(t)$. The output is $B^E = \{b_1', b_2', \ldots, b_N'\}$, where B^E is the envelogram representation of B.

Stockwell transform representation: This was found to be a suitable time– frequency representation for HSs because it avoids the tradeoff between temporal and spectral resolution that afflicts STFT [22], [30]. The S-transform was applied to each signal $b_i(t) \in B$ to produce the S-matrix of the HS, b'. The

output is $B^{TF} = \{b_1', b_2', \ldots, b_N'\}$, where B^{TF} is the time–frequency representation of B.

D. Cluster Analysis

Hierarchical Clustering was applied to both sets of HSs for two purposes.

Outlier rejection: Even sequential S_1 sounds recorded during a short period of time can have large variability [10]. Furthermore, background noise and friction between the recording device and the skin can corrupt signals. To obtain a representation with less variability and fewer outliers, we only considered the segments with the most significant clusters for further analysis. This approach for noise reduction is novel, and our experiments indicate that it is better for the monitoring task described in this paper [29].

Extracting a BL template: This paper introduces features that measure temporal and morphological differences between MS and BL. Signals in the BL set represent the normal cardiac activity of a patient. We wish to extract a patient-specific template representing this activity using compact information. We use the centers of the most significant clusters as the template. The inputs for this stage are the BL and MS sets after transformation to a certain signal representation. We denote these sets as BL and MS. The clustering process is identical for both sets, so we describe it for a general set $B = b_1, b_2, \ldots, b_N$ of N S_1 sounds. We use hierarchical clustering as described previously [23]. The distance between two clusters is defined by

$$d_{GA}(G, H) = \frac{1}{N_G \cdot N_H} \sum_{i \in G} \sum_{j \in H} d_{ij} \qquad (1)$$

where N_G, N_H are the respective number of observations in each group, and d_{ij} represents pairwise observation dissimilarity [31].

The distance metric used to compare observations is the correlation

$$d_{ij} = \|m_i - m_j\| = 1 - \frac{\sum_t (m_{i,t} - \overline{m_i})(m_{j,t} - \overline{m_j})}{\sqrt{\sum_t (m_{i,t} - \overline{m_i})^2} \sqrt{\sum_t (m_{j,t} - \overline{m_j})^2}} \qquad (2)$$

where m_i and m_j are the signals of length n. Additionally, $\overline{m_i} = 1/n \sum_t m_{i,t}$. This distance metric was chosen because it produced more compact clusters than other metrics. To obtain the desired number of clusters, we pruned the hierarchical clustering tree. Observations beneath each cut were assigned to a single cluster.

The cluster analysis assigns a cluster identifier to each signal cycle, using the hierarchical clustering algorithm and dendogram pruning, which results in a clustered set $C = \{(b_1, c_1), (b_2, c_2), \ldots, (b_N, c_N)\}$, where $c_j \in \{1, \ldots, M\}$ are cluster identifiers that are sorted by cluster size, so $|C_1| \geq |C_2| \geq \cdots \geq |C_M|$. Using this notation, a cluster C_j is the set of signal cycles with cluster identifier c_j: $C_j = \{i | (b_i, c_j) \in C\}$. The center of a cluster C_j is the weighted average of the cluster elements, in which each signal cycle is weighted by its similarity to the clusters arithmetic mean

$$\overline{C_j} = \sum_{i \in C_j} w_i b_i, \quad w_i = 1 - D(b_i, (\sum_{i \in C_j} b_i)/|C_j|), \text{ where}$$

D is a distance function [30].

```
 1: procedure TEMPORALSHIFT(B)
 2:    for all S_1 sample b ∈ B do
 3:       for all cluster center C ∈ {C̄₁^BL...C̄ₖ^BL} do
 4:          τ_{b,C} = argmin_τ MSE(C(t), b(t − τ))
 5:          M_{b,C} = MSE(C(t), b(t − τ_{b,C}))
 6:       end for
 7:       C_b = argmin_{C ∈ {C̄₁^BL...C̄ₖ^BL}} {M_{b,C}}
 8:       Set τ_b = τ_{b,C_b}
 9:       Set M_b = M_{b,C_b}
10:    end for
11: end procedure
```

Fig. 5. Temporal shifting algorithm. This algorithm determines the morphologically closest template cluster center, and it determines the temporal delay and morphology change relative to that cluster center. Line 4 computes the temporal shift that minimizes the MSE between the signal and a cluster center. Line 7 finds the cluster center which has the minimum MSE, after temporal shift, with b. The output is used in Sections II-E and II-F.

After M clusters are obtained, the top k largest clusters are denoted as significant clusters, where k is a parameter that needs to be set.

E. Morphology Analysis

The morphology of an S_1 sound encapsulates information about cardiac mechanical activity. The change in morphology from the BL state may correlate with the cardiac stress that the patient experiences (Fig. 2 demonstrates the extent of the morphology change of the first HS). We use this feature as a general indication for change in the physiological state, which is not specific to a certain type of cardiac problem, and may be applicable for identifying numerous cardiac problems that affect cardiac mechanical activity.

The goal is to measure changes in morphology during the ongoing surgery. For this purpose, an MSE estimator is used. The MSE for two vectors with length n is defined as follows:

$$\text{MSE}(x, y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2. \tag{3}$$

The input to this stage is the output of the hierarchical clustering algorithm: The clustered sets of BL and MS. $C_1^{\text{BL}}, \ldots, C_k^{\text{BL}}$ and $\overline{C}_1^{\text{BL}}, \ldots, \overline{C}_k^{\text{BL}}$ are taken to be the k most significant clusters and the cluster centers of BL, respectively. $C_1^{\text{MS}}, \ldots, C_k^{\text{MS}}$ and $\overline{C}_1^{\text{MS}}, \ldots, C_k^{\text{MS}}$ are similarly defined for MS. The cluster center vectors of BL are used as a compact template for the patient's normal cardiac functioning.

To encapsulate the morphology changes alone without considering the temporal shifts, we describe the following temporal shifting algorithm (see Fig. 5). We first define the set of samples we wish to monitor: $B = \{b_i | i \in \bigcup_{j=1}^{k} C_j^{\text{MS}}\}$. It is the set of signals contained in the k most significant clusters of MS.

The temporal shifting algorithm determines the morphological change and temporal delay of each MS signal relative to its closest significant cluster center in the BL set. The average MSE, which is the average change in morphology of MS

samples relative to the BL template, is then calculated

$$\text{MS}_{\text{morph}} = \sum_{b \in \text{MS}} \frac{1}{|\text{MS}|} M_b \tag{4}$$

where M_b is the change in morphology of signal b compared to the template, as defined in the algorithm.

The temporal shifting algorithm can be used when $B = \{b_i | i \in \bigcup_{j=1}^{k} C_j^{\text{BL}}\}$, i.e., when the monitored set is the set of signals from the k most significant clusters of BL. Thus, the average change in morphology of the BL set relative to the patient's template is

$$\text{BL}_{\text{morph}} = \sum_{b \in \text{BL}} \frac{1}{|\text{BL}|} M_b. \tag{5}$$

This value is important because BL may have great variability, which would lead to a less accurate (smeared) representation for the template. This variability represents how BL is different from the template.

The patient's morphology feature, F_{morph}, is defined as follows:

$$F_{\text{morph}} = \ln\left(1 + \frac{\text{MS}_{\text{morph}}}{\text{BL}_{\text{morph}}}\right). \tag{6}$$

This encapsulates the ratio of morphology change between BL and MS on a logarithmic scale. The logarithmic scale was chosen to give less importance to large morphology ratios. Large ratios can occur if BL_{morph} is small.

We note the differences in analysis under the time–frequency signal representation that arise because a signal is represented as a time–frequency matrix. Because of these differences, we use an appropriate MSE measure for the matrices. Additionally, line 4 in the temporal shifting algorithm was changed to allow shifting the signal by frequency f_0. Thus, we minimize the value of $\text{MSE}(C(t, f), b(t - \tau, f - f_0))$.

F. S_1 Onset Analysis

One feature known to be influenced by increased cardiac stress is the time difference in cardiac electromechanical activation (RS_1). Fig. 1 depicts recordings for patients who show significant changes in activation time. We calculate the difference in average RS_1 for the two sets (BL and MS). For this purpose, there is no need to directly calculate the RS_1 timing for the signals. We use the patient's template to calculate, for each signal, the delay in S_1 onset relative to the template. Then, the difference in the average delay between both sets is calculated.

The outputs from the temporal shifting algorithm, when the input is $B = \{b_i | i \in \bigcup_{j=1}^{k} C_j^{\text{MS}}\}$, are the morphological change and temporal delay of each MS signal relative to its closest significant cluster center in BL. The average delay of MS relative to the BL template is, thus, calculated as follows:

$$\text{MS}_{\text{delay}} = \sum_{b \in \text{MS}} \frac{1}{|\text{MS}|} \tau_b. \tag{7}$$

When the input to the algorithm is $B = \{b_i | i \in \bigcup_{j=1}^{k} C_j^{\text{BL}}\}$, the outputs are morphological change and temporal delay for

each signal in BL relative to the template. Thus, the average delay of the BL set relative to its template is calculated as follows:

$$\text{BL}_{\text{delay}} = \sum_{b \in \text{BL}} \frac{1}{|\text{BL}|} \tau_b. \qquad (8)$$

The patient's delay feature, F_{delay}, is defined to be the difference between these two measurements

$$F_{\text{delay}} = \text{MS}_{\text{delay}} - \text{BL}_{\text{delay}}. \qquad (9)$$

To summarize this section, we have described how to produce the morphology feature (F_{morph}) and the delay feature (F_{delay}). The inputs for this process are the sets of S_1 sounds BL and MS, in a specific respiratory phase (inspiration or expiration). We extract a robust patient-specific template to represent normal cardiac activity. The different signal representations that we described (which include time-domain, envelogram, and time–frequency representations) highlight different attributes of the signal. Thus, each signal representation yields different results and corresponds to a different estimator.

G. Monitoring Using Multiple Estimators

Sections II-A to II-F describe the flow of monitoring using a single estimator. The main computational task within this work flow is to estimate the temporal shift between an S_1 sound and the patient's template. For this, we used the approach of shifting the signal until the MSE is minimized between the two signals, as described in the temporal shifting algorithm (see Section II-E). However, for this task, some signal representations are more robust across different patients due to the large variability of S_1 morphology across patients.

For this reason, we introduce the following process to utilize different estimators based on different signal representations. From the three representations, we generate three estimators. We combine the results of these three estimators, or experts, in a "mixture-of-experts" estimator. We accomplish this by taking the average value of the two closest estimators. This produces more robust results across patients.

1) Initial Preprocessing: The signal analysis includes the following stages: 1) preprocessing and segmentation and 2) inspiration and expiration discrimination. These steps are performed for each expert based on the time-domain signal.

At this point, BL and MS, which are in the time domain, are transformed as described in Section II-C to produce a Hilbert envelogram copy (BL^E, MS^E) and an S-transform copy (BL^{TF}, MS^{TF}). Clustering was applied to each of the signal representations separately, as described in Section II-D, to produce the k most significant clusters for each representation. The most significant clusters for the two different signal representations can contain different cardiac cycles. For this reason, the cycles to be monitored were taken to be cardiac cycles that were contained in the k most significant clusters for all three signal representations. The same goes for cycles from the patients BL set.

2) Morphology and S_1 Onset Analysis: The temporal shifting algorithm was applied for every signal representation. The set includes every S_1 sound contained in the k most significant clusters for all signal representations. Each S_1 sound generates three estimates: τ_b^{TD}, τ_b^E, and τ_b^{TF}. These are the temporal shift estimates relative to the patient's template under the time-domain, envelogram, and time–frequency signal representations, respectively. Each representation has its own template.

To utilize the information gained by the three estimators, we define the mixed temporal estimate, τ_b, for sound b to be the average shift for the two most similar estimators

$$\{a, b\} = \text{argmin}_{a,b} |a - b| \text{ s.t. } a, b \in \{\tau_b^{TD}, \tau_b^E, \tau_b^{TF}\}$$
$$\text{and } a \neq b \qquad (10)$$
$$\tau_b = \frac{a + b}{2}. \qquad (11)$$

With this definition, we assume that if two estimators give similar results, they are more likely to be accurate.

After estimating the temporal shift of b relative to its template, we take the morphology change for M_b, to be $M_b = \text{MSE}(C(t), b(t - \tau_b))$, where C is the template cluster center that is morphologically closest to b. This calculation is performed under the time-domain signal representation.

III. RESULTS

A. Experimental Setup

The study group included 15 patients scheduled for elective laparoscopic cholecystectomy, due to symptomatic cholecystolithiasis. The surgeries were performed in Nahariya hospital. The vibro-acoustic HSs were acquired from multiple recording locations. Supplementary data, such as electrocardiogram recordings, were acquired simultaneously.

The patients included six men and nine women, aged between 24 and 92 years. Some patients suffered from heart disease as assessed by medical history, clinical symptoms, ECG, echocardiography, coronary catheterization, and coronary computed tomography. Most heart-diseased patients suffered ischemic heart disease, hypertension, and old myocardial infarction. Cardiac pacing was not included in any of the patients. The details of the medical data are included in Table I. From here on, patients who suffered from a heart disease are referenced as "ill," while other patients without any heart diseases are referenced as "healthy."

Surgery was performed under general anesthesia and positive pressure pneumoperitoneum was established by CO_2 gas that was kept automatically under pressure of 14 mmHg. Following introduction of a video camera and additional canulae, the gallbladder was dissected, resected, and pulled out of the abdominal cavity by an endobag. Afterwards, the gas was removed to desuflate the abdominal cavity, before completion of the operation.

Recording was performed during three different phases of the surgery: After anesthesia, during pneumoperitoneum, and after desufflation. Each recording took at least 30 s. The recordings were labeled according to operation stages. The recorded signals were saved digitally to the hard disk of the recording system.

TABLE I
DETAILS OF LAPAROSCOPIC PATIENTS

ID	Age	Gender	Diagnosis	Details For Cardiac Diseases
AZ	29	F	Biliary colic Cholelithiasis Obstructive jaundice	
AJ	24	F	Cholelithiasis	
KA	69	M	Acute pancreatitis, Calculus of Gallbladder	I.H.D, old M.I, LVH, s/p CVA
DM	84	M	Acute pancreatitis, Calculus of Gallbladder	Old MI, I.H.D, hypertension, type II D.M, CRF, s/p CVA, LVEF 48%
RB	60	F	Cholelithiasis, Calculus of Gallbladder	
SHA	43	M	Calculus of Gallbladder	IHD, SVT, hypertension, obesity
PZ	92	F	Acute pancreatitis, Calculus of Gallbladder Cholelithiasis	
YY	86	M	Calculus of Gallbladder Obstructive jaundice	COPD, IHD, DM, moderate MR,
PA	76	F	Acute pancreatitis, Calculus of Gallbladder	
ES	35	F	Choledocholithiasis Cholelithiasis, Calculus of Gallbladder Obstructive jaundice	Hypothyroidism, effort dyspnea
NA	27	F	Cholelithiasis, Calculus of Gallbladder	
DE	53	F	Calculus of Gallbladder	
SUA	29	F	Cholelithiasis, Calculus of Gallbladder	
GR	58	M	Obstructive jaundice Calculus of Gallbladder	IHD, s/p MI (infero posterior and Rt, including stent), DM, akinetic basal, septal and dorsal wall on echocardiography
MM	28	M	Cholelithiasis, Calculus of Gallbladder	Hypertension, DM, clinical presentation of CHF

TABLE II
MEAN NUMBER OF PROCESSED CARDIAC CYCLES ACROSS SUBJECTS

Phase	Inspiration	Expiration
Before anesthesia	21.3 ± 8.3	19.8 ± 11.7
During insufflation	28.4 ± 21.4	16.6 ± 10.1
After desufflation	18.1 ± 6.7	14.6 ± 6.1

B. Parameter Selection

The physiological monitoring framework was applied individually to data for each of the 15 patients. The data were constructed from a single recording channel, chosen based on its recording quality. Only the S_1 components of the HS cycles were used. The mean number of cardiac cycles processed across subjects, per surgery phase (before induction of anesthesia, during abdominal insufflation or right after desufflation) and per respiration phase (inspiration or expiration) is detailed in Table II.

After signal preprocessing, the signals were transformed into three signal representations: Time-domain, Hilbert envelogram, and the S-transform (time–frequency representation). Analysis of the resulting time–frequency matrix indicates that most of the energy is concentrated at frequencies between 10 and 40 Hz. For this reason, we take 100 nonoverlapping frequency bands in this frequency range for the S-transform analysis.

Hierarchical clustering was applied to the dataset, and the clustering tree obtained was pruned to obtain four clusters. The number of clusters was empirically tuned; experiments with fewer clusters showed results that were not sensitive enough for different morphologies. Experiments with more than four clusters resulted in sparse clusters. The number of significant clusters was set to $k = 2$. This means that the data in the two largest clusters were considered for further analysis, and the data in the two smallest clusters were considered outliers. On average, 76.8% of the data were contained in the two largest clusters. For the mixture of estimators, only signals contained

in the two largest clusters of all signal representations were analyzed. This intersection covers 67.8% of the original data, on average. This indicates that the clustering processes for different signal representations resulted in similar dendograms.

C. Qualities of Estimators

Our framework extracts features from HSs. In the process of extraction, we estimate the relative delay, τ_b, in the onset between an S_1 sound, b, and a template, T, in one of the signal representations. This estimation is performed by using the temporal shifting algorithm, i.e., by shifting b until a minimal MSE is obtained between the two signals. This is denoted by

$$\hat{\Theta}(b, T) = \tau_b. \tag{12}$$

For an unbiased estimator, the error is $\mathrm{MSE}(\hat{\Theta}) = Var(\hat{\Theta})$ [32]. We compared the qualities of our estimators by calculating the average sample standard deviation of τ_b across patients, which is equal to the estimator error, for each phase of surgery and for each estimator. Fig. 6 indicates that for all phases of surgery, the errors of the mixture-of-experts estimator and the raw signal estimator are the lowest.

D. Estimator Monitoring Results

The results for the extracted features for each patient and for each phase of surgery are presented. The results are presented only for the mixture-of-experts estimator (see Fig. 7) because it was found to be best in terms of error rate and manual assessment. Full results can be found in another work [29]. The results are presented only for the inspiration respiratory phase. The reason for this is that the artificial inspiration induces more stress than artificial expiration.

We see that for the introduced estimator, there is a shift in the extracted features relative to the BL state. It is evident that patients who are tagged as having cardiac problems experience a more severe shift in their extracted features during insufflation. We notice that this estimator can be used for a high-quality
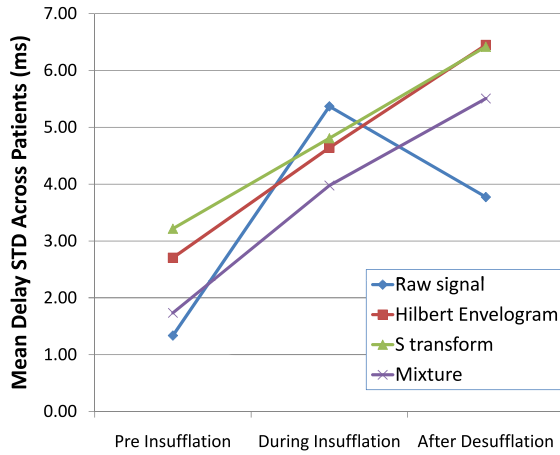
Fig. 6.   Mean sample STD across patients for each phase of surgery and for each type of estimator.
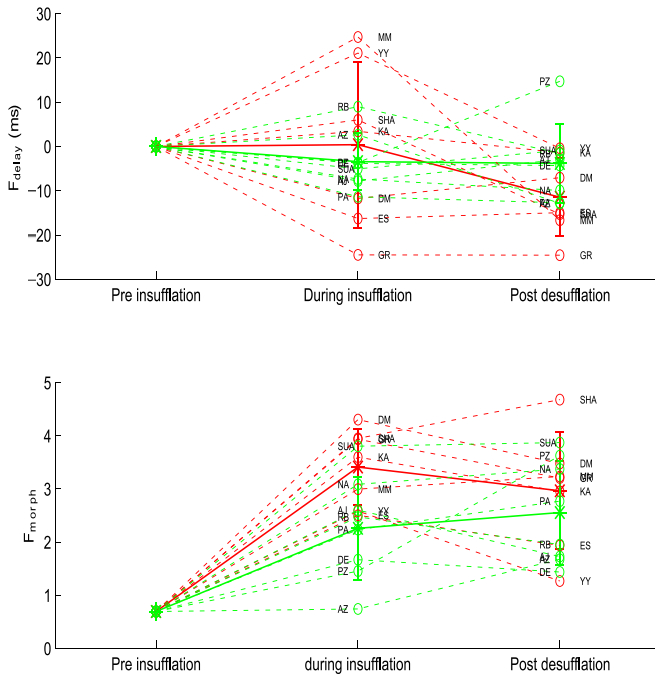

Fig. 7.   Mixture-of-experts estimator output for the relative delay feature (top) and relative morphology feature (bottom) for each of the 15 patients during the three different surgery phases. Green circles represent healthy patients, and red circles represent ill patients. Initials also appear for each patient. A line represents the temporal changes for a given patient. The mean for each group of patients is represented by a star. Error bars represent the STD for each group.

classification of ill and healthy patients during insufflation. It is also evident that right after abdominal desufflation, the ability to discriminate between ill and healthy patients deteriorates. Moreover, for some of the patients, it is clear that the value for the extracted features gets closer to the BL. These findings indicate that the features we have chosen are beneficial for monitoring cardiac stress.

TABLE III
CLASSIFICATION ACCURACY FOR EACH FEATURE,
MONITORED PHASE, AND ESTIMATOR

Feature	Phase	Raw Signal	Hilbert Transform	S Transform	Mixture
F_{delay}	During	0.6	0.87	0.67	0.73
F_{morph}	insufflation	0.73	0.33	0.4	0.67
Both		0.8	0.8	0.8	0.8
F_{delay}	After	0.73	0.07	0	0.47
F_{morph}	desufflation	0.4	0.4	0.4	0.33
Both		0.73	0.13	0.33	0.47

TABLE IV
P-VALUE SCORE OF WELCH'S T-TEST BETWEEN HEALTHY AND ILL PATIENTS
FOR EACH FEATURE, MONITORED PHASE, AND ESTIMATOR

Feature	Phase	Raw Signal	Hilbert Transform	S Transform	Mixture
F_{delay}	During	0.15	0.05	0.02	0.03
F_{morph}	insufflation	0.17	0.42	0.4	0.02
F_{delay}	After	0.06	0.96	0.89	0.33
F_{morph}	desufflation	0.62	0.59	0.47	0.46

E. Healthy and Ill Patient Discrimination

1) Classification: The presented framework extracts the relative delay and relative morphology features for each monitoring phase. Section III-D presents the evolution of both of the features during different phases of surgery for all patients. Looking at this graph, it appears that the presented features can be used to discriminate between ill and healthy patients during insufflation. We tested this assumption by examining the following classification problem: For each monitoring phase, feature, and estimator, we calculated the accuracy of classifying patients as healthy or ill. The clinical condition of each patient is detailed in Section III-A. We used an SVM classifier with a linear kernel. Because the number of samples is small (only 15 patients), we used leave-one-out cross validation.

Table III indicates that classification using both of the features gave good results during insufflation (80% accuracy). The average accuracy (for F_{delay}, F_{morph}, and for both features) is best with the mixture estimator ($73 \pm 6\%$ accuracy). This agrees with previous results presented in Sections III-C and III-D.

We also note that after desufflation, the classification quality using the detailed features becomes low, even less than that expected by chance. This is reasonable because in this phase of surgery, there is no abdominal insufflation, and the cardiac function of the patient becomes more similar to the BL.

2) Significance Testing: We tested whether the distribution of healthy patients is similar to that of ill patients. We used Welch's t-test and applied it to the same data used in the classification problem. Table IV suggests that the only estimator that is significant for both features during insufflation is the mixture-of-experts estimator. This result matches the results presented earlier regarding classification accuracy for this classifier, and it is in agreement with error of this estimator and our manual assessment. The Hilbert envelogram and S-transform estimator were also significant for F_{delay}.

TABLE V
CLASSIFICATION ACCURACY FOR EACH ECG FEATURE
AND MONITORED PHASE

Feature	Phase	ECG Raw Signal
F_{amp}	During	0.4
F_{morph}	insufflation	0.6
Both		0.6
F_{amp}	After	0.4
F_{morph}	desufflation	0.47
Both		0.53

We also note that after desufflation, there is no significant difference for all features and all estimators. This matches the classification results obtained in the last section.

F. ECG Monitoring Comparison

We further examined two ECG features as alternatives to the delay and morphology features extracted from HS. These features were the ECG morphology and the power ratio of the R wave amplitude between the BL phase and the monitored phase. Patient discrimination based on these ECG features was compared to the HS feature quality.

ECG signals were recorded at the same time as HSs. After preprocessing the HSs and separating them into inspiration or expiration phases, we analyzed the corresponding ECG signals. We use the same notation as for HSs; therefore, BL is the set of ECG signals recorded during the BL phase, and MS is the set of ECG signals recorded during the monitored phase. Only the raw signal representation was examined. Clustering was performed for the BL set with the same parameters as for HSs to produce a template for each patient (consisting of the two most significant cluster centers).

The ECG morphology feature was extracted in the same way as described in Section II-E. The power ratio feature was calculated as the ratio between the mean R wave amplitude of the signals in MS and BL. Formally, for an ECG signal b, we define $A(b)$ as the amplitude of the R wave of b. Then, F_{amp} is defined as follows:

$$F_{\mathrm{amp}} = \frac{\frac{1}{|\mathrm{MS}|} \sum_{b_i \in \mathrm{MS}} A(b_i)}{\frac{1}{|\mathrm{BL}|} \sum_{b_i \in \mathrm{BL}} A(b_i)}. \tag{13}$$

Table V and Fig. 8 indicate that HS features were better than ECG features for discriminating between sick and healthy patients during insufflation. Of the two ECG features, the morphology feature produced better discrimination. Table VI suggests that there is no significant difference between the distributions of sick and healthy patients for any of the features. The morphology feature performed better here as well.

These findings indicate that overall the ECG features are not effective for monitoring cardiac stress and that monitoring using the acoustic S_1 sounds, which indicate the mechanical activity of the heart, provides useful additional information that cannot be obtained from the ECG signal alone.
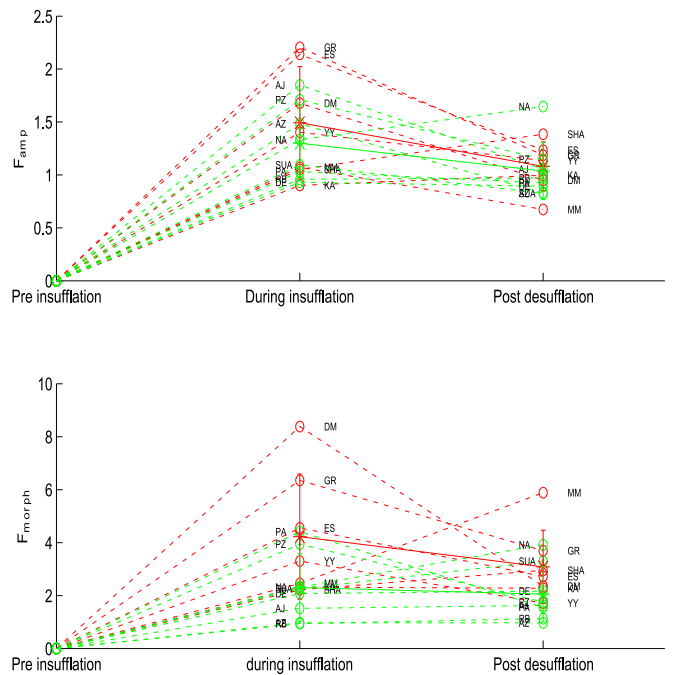

Fig. 8. ECG raw signal estimator output for the power ratio feature (top) and the relative morphology feature (bottom) for each of the 15 patients during the three different surgery phases. Green circles represent healthy patients, and red circles represent ill patients. Initials also appear for each patient. A line represents the temporal changes for a given patient. The mean for each group of patients is represented by a star. Error bars represent the STD for each group.

TABLE VI
P-VALUE SCORE OF WELCH'S T-TEST BETWEEN HEALTHY AND ILL PATIENTS
FOR EACH ECG FEATURE AND MONITORED PHASE

Feature	Phase	ECG Raw Signal
F_{amp}	During	0.4
F_{morph}	insufflation	0.6
F_{amp}	After	0.4
F_{morph}	desufflation	0.47

IV. CONCLUSION

This paper studied a well-known problem of increased cardiac risk during noncardiac surgery. We proposed a novel monitoring framework based on the mechanical activity of the heart, which is represented by the sounds emitted from the heart. One novel feature of our method is the use of a nonparametric measure of morphology changes between the individual BL morphology and the continuous measure of cardiac electric conductance time. The other novelty is the use of clustering to remove outliers and to ensure robustness against noise during the HS recording, which is performed in a noisy surgery room. The third novelty is in the computational technique used to measure the delay in S_1 onset relative to a different signal.

We examined various types of estimators that use different signal representations for the extraction of these measures. Additionally, we described a mixture-of-experts estimator, which utilizes measurements of different estimators to produce a more

robust measurement. We have demonstrated an improvement in accuracy and robustness using this mixture-of-experts approach.

We have demonstrated that the two physiological measures (morphology change and delay in S_1 onset relative to BL) undergo changes during the cardiac stress caused by abdominal insufflation. We have further shown that shortly after desufflation, the above features start returning to their BL values. We have also shown that for patients with known cardiac problems, the proposed physiological features change more dramatically, and in some cases, they start returning to their BL values more slowly. Furthermore, setting a threshold for the values of the extracted features generated two groups of patients. We have shown that these groups are similar (with 80% accuracy) to the sick and healthy groups of patients, which were labeled according to their cardiac state prior to surgery.

We have further shown here and elsewhere that the mechanical functional changes during cardiac stress are not as readily manifested in ECG measurements, such as ECG morphology, suggesting that our proposed features, which are based on HSs, add new information to the monitoring of cardiac functionality under cardiac stress caused during surgery. These HS measurements may also be useful in other cases, such as the onset of congestive heart failure or before dialysis in kidney failure patients.

REFERENCES

[1] S. Hoeks *et al.*, "Cardiovascular risk assessment of the diabetic patient undergoing major noncardiac surgery," *Best Practice Res. Clin. Endocrinol. Metabolism*, vol. 23, no. 3, pp. 361–373, 2009.

[2] D. T. Mangano, "Perioperative medicine: NHLBI working group deliberations and recommendations," *J. Cardiothoracic Vascular Anesthesia*, vol. 18, no. 1, pp. 1–6, 2004.

[3] P. Devereaux *et al.*, "Perioperative cardiac events in patients undergoing noncardiac surgery: A review of the magnitude of the problem, the pathophysiology of the events and methods to estimate and communicate risk," *Can. Med. Assoc. J.*, vol. 173, no. 6, pp. 627–634, 2005.

[4] P. Devereaux *et al.*, "Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (poise trial): A randomised controlled trial," *Lancet*, vol. 371, no. 9627, pp. 1839–1847, 2008.

[5] A. Biagini *et al.*, "Unreliability of conventional visual electrocardiographic monitoring for detection of transient st segment changes in a coronary care unit," *Eur. Heart J.*, vol. 5, no. 10, pp. 784–791, 1984.

[6] J. M. Leung *et al.*, "Prognostic importance of postbypass regional wall-motion abnormalities in patients undergoing coronary artery bypass graft surgery," *Anesthesiology*, vol. 71, no. 1, pp. 16–25, 1989.

[7] J. M. Leung *et al.*, "Automated electrocardiograph ST segment trending monitors: Accuracy in detecting myocardial ischemia," *Anesthesia Analgesia*, vol. 87, no. 1, pp. 4–10, 1998.

[8] G. Amit *et al.*, "Automatic analysis of vibro-acoustic heart signals." *Tel-aviv University*, Jul. 2008.

[9] T. R. Reed *et al.*, "Heart sound analysis for symptom detection and computer-aided diagnosis," *Simulation Model. Practice, Theory*, vol. 12, no. 2, pp. 129–146, 2004.

[10] G. Amit *et al.*, "Respiratory modulation of heart sound morphology," *Amer. J. Physiology-Heart Circulatory Physiol.*, vol. 296, no. 3, pp. H796–H805, 2009.

[11] D. Flores-Tapia *et al.*, "Heart sound cancellation based on multiscale products and linear prediction," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 2, pp. 234–243, Feb. 2007.

[12] S. R. Messer *et al.*, "Optimal wavelet denoising for phonocardiograms," *Microelectronics J.*, vol. 32, no. 12, pp. 931–941, 2001.

[13] F. Ghaderi *et al.*, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 12, pp. 3360–3367, Dec. 2011.

[14] Z. Yan *et al.*, "The moment segmentation analysis of heart sound pattern," *Comput. Methods Prog. Biomed.*, vol. 98, no. 2, pp. 140–150, 2010.

[15] C. Ahlstrom *et al.*, "A method for accurate localization of the first heart sound and possible applications," *Physiol. Meas.*, vol. 29, no. 3, pp. 417–428, 2008.

[16] D. Gill *et al.*, "Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model," in *Proc. Comput. Cardiol.*, 2005, pp. 957–960.

[17] C. N. Gupta *et al.*, "Neural network classification of homomorphic segmented heart sounds," *Appl. Soft Comput.*, vol. 7, no. 1, pp. 286–297, 2007.

[18] T. lmez and Z. Dokur, "Classification of heart sounds using an artificial neural network," *Pattern Recog. Lett.*, vol. 24, no. 1, pp. 617–629, 2003.

[19] Z. Dokur and T. lmez, "Heart sound classification using wavelet transform and incremental self-organizing map," *Digital Signal Process.*, vol. 18, no. 6, pp. 951–959, 2008.

[20] S. Choi and Z. Jiang, "Comparison of envelope extraction algorithms for cardiac sound signal segmentation," *Expert Syst. Appl.*, vol. 34, no. 2, pp. 1056–1069, 2008.

[21] A. Moukadem *et al.*, "Phonocardiogram signal processing module for auto-diagnosis and telemedicine applications," 2012, [Online]. Available: http://www.intechopen.com/books/authors/ehealth-and-remote-monitoring/phonocardiogram-signal-processing-module-for-auto-diagnosis-and-telemedicine-applications.

[22] G. Livanos *et al.*, "Heart sound analysis using the S transform," in *Proc. Comput. Cardiol.*, 2000, pp. 587–590.

[23] S. Kofman *et al.*, "Discovery of multiple level heart-sound morphological variability resulting from changes in physiological states," *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 315–324, 2012.

[24] A. Yadollahi and Z. M. Moussavi, "A robust method for heart sounds localization using lung sounds entropy," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 497–502, Mar. 2006.

[25] J. L. Joris *et al.*, "Hemodynamic changes during laparoscopic cholecystectomy," *Anesthesia Analgesia*, vol. 76, no. 5, pp. 1067–1071, 1993.

[26] A. Leatham, "Splitting of the first and second heart sounds," *Lancet*, vol. 264, no. 6839, pp. 607–614, 1954.

[27] S. Efstratiadis and A. D. Michaels, "Computerized acoustic cardiographic electromechanical activation time correlates with invasive and echocardiographic parameters of left ventricular contractility," *J. Cardiac Failure*, vol. 14, no. 7, pp. 577–582, 2008.

[28] N. de Oliveira Neto *et al.*, "Abnormalities of the systolic time intervals obtained by electronic stethoscope in heart failure," *Internet J. Cardiol.*, vol. 5, no. 2, 2008, [Online]. Available: https://ispub.com/IJC/5/2/3503

[29] J. Herzig, "Monitoring cardiac stress using features extracted from S1 heart sounds," Masters dissertation, Sch. of Comput. Sci., Tel-Aviv Univ., 2014.

[30] G. Amit *et al.*, "Cluster analysis and classification of heart sounds," *Biomed. Signal Process. Control*, vol. 4, no. 1, pp. 26–36, 2009.

[31] T. Hastie *et al.*, *The Elements of Statistical Learning*, vol. 1. New York, NY, USA: Springer-Verlag, 2001.

[32] S. M. Kay, *Fundamentals of Statistical Signal Processing, I: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.