Design and Comparative Performance of a Robust Lung Auscultation System for Noisy Clinical Settings

Ian McLane, Dimitra Emmanouilidou, James E West, and Mounya Elhilali

***Abstract*—Chest auscultation is a widely used clinical tool for respiratory disease detection. The stethoscope has undergone a number of transformative enhancements since its invention, including the introduction of electronic systems in the last two decades. Nevertheless, stethoscopes remain riddled with a number of issues that limit their signal quality and diagnostic capability, rendering both traditional and electronic stethoscopes unusable in noisy or non-traditional environments (e.g., emergency rooms, rural clinics, ambulatory vehicles). This work outlines the design and validation of an advanced electronic stethoscope that dramatically reduces external noise contamination through hardware redesign and real-time, dynamic signal processing. The proposed system takes advantage of an acoustic sensor array, an external facing microphone, and on-board processing to perform adaptive noise suppression. The proposed system is objectively compared to six commercially-available acoustic and electronic devices in varying levels of simulated noisy clinical settings and quantified using two metrics that reflect perceptual audibility and statistical similarity, normalized covariance measure (NCM) and magnitude squared coherence (MSC). The analyses highlight the major limitations of current stethoscopes and the significant improvements the proposed system makes in challenging settings by minimizing both distortion of lung sounds and contamination by ambient noise.**

***Index Terms*—Stethoscope, biomedical acoustics, adaptive signal processing, signal to noise ratio, public healthcare.**

I. Introduction

RESPIRATORY diseases and illnesses pose some of the most serious threats to individuals and to the public, affecting over 600 million people worldwide; early and accurate diagnosis of respiratory conditions presents a major challenge toward successful triage and treatment [1]. The use of chest auscultation to rapidly screen for and diagnose lung diseases and infections has been in practice since the invention of the stethoscope by R.T.H. Laennec in 1816 [2]. Stethoscopes are used in respiratory examination to distinguishing normal respiratory sounds from crackles, wheezes, and pleural rub in order to make correct diagnoses and remain a key diagnostic device due to their portability, low cost, and noninvasive nature.

Despite its universal adoption, the use of the stethoscope is riddled by a number of issues including subjectivity in interpretation of chest sounds, interlistener variability and inconsistency, need for medical expertise, and vulnerability to airborne ambient noise. The presence of background noise is particularly important because acoustic systems have several noise transmission paths–chestpiece, tubing, and earpieces–through which airborne noise tends to 'leak' and mask body sounds, rendering the interpretation of these sounds challenging or impossible and exacerbating the other issues related to subjectivity and variability. In global health or emergency medicine environments, noise levels far exceed those of a clinical environment (approximately 75 dB SPL) and frequently exceed the approximate upper limit of correct detection of abnormal breath sounds (an average of 81 dB SPL) [3]. Health care providers in resource-limited or non-traditional settings rely upon low-cost tools or clinical presentations to make critical patient management decisions, rarely relying on auscultation due to these limitations [4]. However, one study showed that including auscultatory findings in the WHO case management guidelines for pneumonia significantly increased correct identification of radiographically-confirmed pneumonia by 16% [5].

Electronic stethoscopes were introduced in the last few decades, opening up new opportunities for sound amplification, enhanced frequency range, quality improvement, and computerized analysis. Passive filtering improves the sound quality, but these systems are still susceptible to dynamic noise seen in most real-world environments [6]. In many settings, background chatter and other environmental noises are common, and patient motion contaminates the sound signal picked up by the stethoscope. Several electronic systems (e.g., Littmann 3200 and Eko Core) are additionally plagued by the same airborne sound transmission paths as acoustic stethoscopes by using a traditional chestpiece or a small loudspeaker and traditional binaural tubing.

Manuscript received April 7, 2020; revised September 14, 2020 and November 23, 2020; accepted January 24, 2021. Date of publication February 3, 2021; date of current version July 20, 2021. This work was supported in part by the National Institutes of Health under Grants R01HL133043, U01AG058532, and R43MD014104 and in part by the Office of Naval Research under Grants N000141912014 and N000141912689. *(Corresponding author: Mounya Elhilali.)*

The authors are with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: imclane1@jhu.edu; dimitra.emmanouilidou@gmail.com; jimwest@jhu.edu; mounya@jhu.edu).

Digital Object Identifier 10.1109/JBHI.2021.3056916

Therefore, most acoustic and electronic commercial devices are meant to be used only in controlled clinical environments (e.g., a quiet doctor's office) or in stationary noise. They tend to be unusable in noisy clinical settings such as emergency or operating rooms, ambulances, or outpatient or rural clinics. Because of these limitations, development of automated approaches based on computer-aided auscultation for these electronic systems also remain in their infancy [7]–[10].

The aim of this work is to design an adaptable and multipurpose device that could be used across the spectrum of clinical environments by any type of healthcare provider, and was uniquely designed to overcome known limitations of current devices. This paper presents a comprehensive platform which includes both hardware and software redesign to improve sensing, acquisition, and analysis of body-generated sounds. The proposed system comprises a collection of sensing and functional components—a microphone array, an external facing microphone, an audio codec, and on-board processing—packaged as a standalone device that can be deployed into the field. The microphone array maximizes the measurement sensitivity and uniformity, while the external facing microphone is used for adaptive noise suppression. The algorithm was designed specifically for auscultation, dynamically adapting to the environment and actively suppressing unwanted interfering noise, including noises with a spectral profile and temporal signature that overlap with the body sounds.

In order to objectively quantify the design improvements, the proposed device is compared against 6 commercially available stethoscopes ranging from acoustic designs to electronic state-of-the-art technologies. These specific devices were selected because they each represent unique construction and parameters that influence the characteristics of the transmitted sound and deliver a different performance. The performance of the systems is accounted for in terms of several key metrics: 1) sensitivity and uniformity of the transducing mechanism, 2) agreement with and quality-preservation of the transmitted body signals, and 3) resistance to unwanted ambient noise events.

The type of validation that is presented in this paper is not reasonable to do in real-world scenarios with this number of devices: exact comparison is limited by the reproducibility of body sounds in patients and ambient noise profiles in the real world. We therefore use custom setups for reproducibility. For sensing sensitivity, an 9-driver array was used to determine position-dependency and sensitivity of the sensor. For quality-preservation of body sounds and noise robustness, a chest sound simulator to emit signals recorded from normal and abnormal patients is placed in an artificial clinical setting, created using a sound booth wherein an 6-speaker setup broadcasts noises of various types and volume levels. A combination of different objective quality measures adapted from speech processing are used to report performance for the various systems. By using these setups, we are able to directly compare the proposed device to many devices in hundreds of combinations of lung sounds and ambient noises, and eliminate uncertainty across the analysis.

The sensing sensitivity analysis reveals that acoustic systems with a diaphragm are highly sensitive to placement, while the proposed system provides uniformity across the entire pick-up surface. The noise robustness results reveal that there is a mix in the performance of auscultation systems: some perform best in quiet conditions while others prove effective in noisy environments. The proposed hardware system with real-time adaptive noise suppression maintains a low coherence with ambient noise while preserving the lung sounds as much as possible, outperforming the commercial devices in all simulated environments.

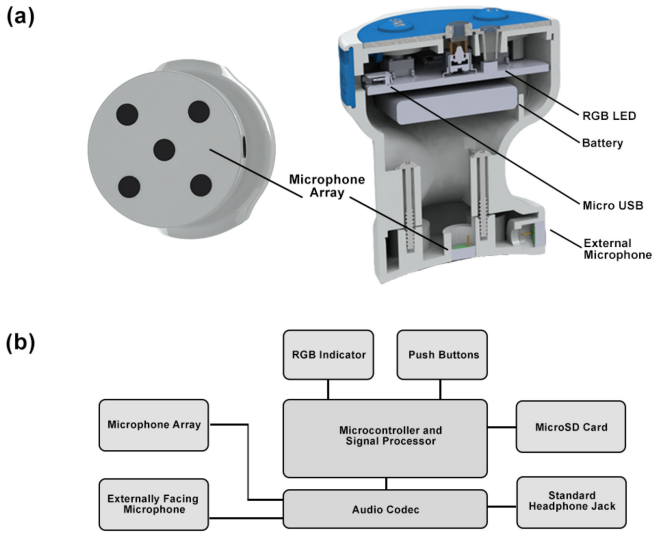

Fig. 1. Illustrations and photo the device. (a) Three-dimensional rendering of the proposed smart stethoscope slice to demonstrate the internal hardware components. (b) Basic circuit diagram and operation of the proposed system.

II. System Architecture

The proposed programmable stethoscope (called the JHUscope here) mitigates a number of limitations in existing auscultation systems with an on-board microprocessor and a unique sensor design. The sensor design increases sensitivity while the microprocessor enables real-time capabilities such as adaptive noise suppression and automatic volume control. The design also incorporates other features including on-board memory, battery life optimization, and patient motion detection. Of primary importance, it was designed to be small, dependable, and simple to operate. Overall dimensions are 40 mm (height) $\times$40 mm (diameter), and 45.5 g in weight.

A. Hardware

The core of the digital hardware design includes a microphone array, an external facing microphone, an audio codec, a microprocessor, and a headphone jack for listening (Fig. 1). The acoustic array consists of five omnidirectional electret microphones (Shenzhen Horn Electroacoustic Technology, China), in addition to an externally-facing microphone that is matched with the array microphones and used for adaptive noise suppression. The patient-facing microphones are laid out in a cross pattern to maximize uniform surface sensitivity, a topic we examine further in Section III-A. The audio signals from the sensing array are summed with identical and time-invariant weight factors, and

both the array signal and ambient signal are processed through a 96 kHz, 24-Bit ultra-low power NXP audio codec (NXP Semiconductors, Netherlands), which coordinates data acquisition, analog-to-digital conversion, and transfer to the digital signal processing (DSP) block of the microprocessor.

The microprocessor is a Kinetis microcontroller with an ARM Cortex-M4 core (NXP Semiconductors, Netherlands). The Kinetis microprocessor was chosen because of its small footprint, low power consumption, and low-cost without compromising processing speed and DSP-specific architecture.The embedded microprocessor is a key component of the JHUscope hardware, as it aims to run a number of algorithms to enable more intelligent capabilities to the system. In its current form, the microcontroller is used to implement the real-time processing discussed in Section II-B. The inclusion of the microcontroller also allows for increased flexibility in future iterations, including the addition of other sensors such as inertial measurements to augment analysis methods.

Because this system is intended to be used in clinical, emergency medical, and global health environments, it was also carefully designed in a way that offers optimal battery consumption. The device is powered by 3.7 V/500 mAh Li+ battery. The system has an average current draw of 75 mA with peak current draws up to 120 mA and sleeping current draw less than 3 mA, meaning the battery lasts close to 6 hours of continuous usage or 2-3 days of normal clinical usage.

In addition, the design includes on-board storage for data collection in the field and a micro-USB interface for recharging the battery of the device. The device is controlled through push buttons to control volume, initiate data recording, and powering the device on and off. LED indicators serve as simple visual guides for the user during the operation of the system. Users can listen in real-time through a standard 3.5 mm headphone jack, which is enabled through the same NXP audio codec used for signal acquisition (Fig. 1).

The device can also be easily reprogrammed through the micro-USB connected to a computer. Due to the programmable nature of the JHUscope, this same hardware system can be programmed to implement updates to signal processing, usability features, and real-time computerized lung sound analysis (CLSA). Noise suppression combined with CLSA would enable this system to be deployed in all types of traditional and non-traditional clinical settings to provide objective and repeatable respiratory assessment at a low cost.

B. Software

One of the key challenges of performing noise cancellation in auscultation systems is to deliver a noise-free signal that completely removes any distortions from airborne ambient noise, all while cautiously avoiding cancellation of auscultation patterns that are indicative of disease and can easily masquerade as noise (e.g., crackling noises from the lung masquerading as patient-generated noise).

The simple addition of active noise cancelling (ANC) headphones to an electronic stethoscope does not adequately address these challenges because ambient noise is only controlled at one entry point (the earpiece), but does not control for airborne ambient noise contamination at the chestpiece. Therefore, the JHUscope runs a real-time adaptive noise suppression technique on the device that optimizes both the removal of ambient noise and the preservation of the signal in question, with better quality than existing systems. At the core, the algorithm is an extension of classic spectral subtraction used in the fields of communication and speech enhancement in that it operates in the frequency domain and suppresses the ambient noise spectrum $d(n)$ from the noisy auscultation spectrum $y(n)$ [11], [12]. A full account and validation of the algorithm has been presented before (see [13] for complete details), but the real-time implementation will be summarized next.

The device captures both $y(n)$ and $d(n)$ simultaneously through a dual microphone setup, as described in Section II-A. The algorithm then maps short-time frames of the two signals into the spectral domain using a short-term Fourier transform, resulting in signals $Y_\tau(\omega_k)$ and $D_\tau(\omega_k)$; where ω_k indexes $k = 1, \ldots, 32$ frequency bands and τ is used to represent processing over short-time windows $w(n)$. The ambient noise is then suppressed in an adaptive manner, such that

$$|\hat{X}_\tau(\omega_k)|^2 = |Y_\tau(\omega_k)|^2 - \gamma_\tau \lambda_k |D_\tau(\omega_k)|^2 \quad (1)$$

where γ_τ and λ_k are time and frequency scaling coefficients, respectively. These coefficients are adjusted automatically for each 50 ms time frame τ by the current SNR of that time frame, and for each frequency band k by the spectral profile of the signals and a priori knowledge of body sound profiles [14]. An estimate of the true lung sound signal $\hat{x}_\tau(t)$ is then obtained by applying the inverse Fourier transform on $\hat{X}_\tau(\omega_k)$ and then implementing a real-time version of overlap-and-add using circular buffers. This localized time and frequency treatment is especially crucial given the variable, unpredictable, and nonuniform nature of noise that overlap in time and in frequency with lung sounds.

This algorithm was previously tested on real pediatric data collected from busy or remote clinic centers with challenging noise environments, part of a large-scale pneumonia etiology study (PERCH) [15]. Lung signals from 22 infants acquired in Gambia were analyzed using this algorithm. Objective quality measures were used to compare the algorithm to an Adaptive-Noise-Canceling scheme (FXLMS) [16], [17]. While the FXLMS algorithm was shown to deliver reasonable signal quality under controlled noise environments, it was unable to adapt fast enough to the high non-stationary nature of noise in a real-life, busy clinic. Furthermore, formal listening tests were performed by presenting a panel of 17 experts, a majority of whom were pediatric pulmonologists, with auscultation signals with and without processing through the proposed denoising algorithm in a blind listening judgment of signal quality. An overwhelming 95.1% of cases processed through the proposed algorithm were preferred by the panel of experts, further confirming the efficacy of the denoising software at delivering improved quality signals without removing the clinically-important information in the auscultation signal [13].

Finally, the system also includes an an automatic volume control (AVC) block to amplify signals of interest and reduces loud, explosive signals caused by device or patient motion or

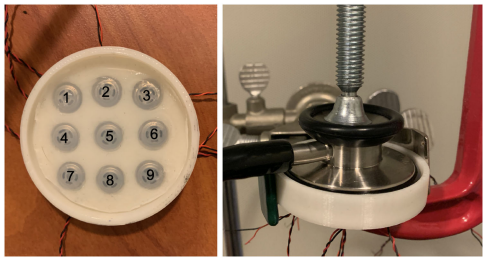

Fig. 2. Nine driver array setup (numbered) to determine surface sensitivity (left) and an example of how devices were affixed to the driver array during recording (right).

friction. The AVC is designed as a traditional audio compressor that decreases the signal when the measured level is above the programmed threshold. The threshold, the attack rate (the rate at which the incoming signal is attenuated down to the threshold) and the decay rate (the rate at which this volume returns back to its original level) are programmable, and customized so that the AVC does not distort important clinical sounds such as heart sounds.

III. Methods

Two specific performance measures were chosen to compare the JHUscope to the commercially-available systems: the JHUscope's performance needs to balance high sensitivity with noise robustness to ensure that the device can be used by all personnel in any type of environment. The following setups were used to perform objective and repeatable assessment based on these performance measures.

A. Sensing Sensitivity

Sensing sensitivity examines the optimal positioning against the body by determining the sensitivity of the system's pickup surface. The sensing sensitivity is assessed by clamping the head of the JHUscope and subsequently, for comparison, a representative acoustic stethoscope (ADC Adscope), and a representative electronic stethoscope (Thinklabs One) directly on top of a 3-by-3 array of 8.5mm-diameter headphone drivers (Shenzhen Horn Electroacoustic Technology, China) spaced 4.5 mm apart and covered with a 2.5 mm layer of Ecoflex (Smooth On Inc., PA, USA) to account for the inertial loading introduced by the skin on the diaphragm. Ecoflex was chosen because of the similar mechanical properties to that of human skin [18]. See Fig. 2 for a depiction of the driver array setup.

A low-amplitude driving signal of pink noise $x_p(t)$ is played from each driver position successively and output signals $y_{i_j}(t)$ are recorded for the individual driver positions $i = [1, .., 9]$ over $j = [1, .., 20]$ trials. Two comparisons are done to assess the role location and placement plays in the acoustic and proposed auscultation systems. The comparisons are only done with the diaphragm side of the acoustic stethoscope due to findings that deny the claims there is a significant difference in filtering between the diaphragm and bell settings [19].

TABLE I
Speaker Placement Relative to the Position of the Stethoscope on the Chest Sound Simulator

Speaker:	**1**	**2**	**3**	**4**	**5**	**6**
h (m)	1.55	0.88	1.50	0.76	0.00	0.77
d (m)	1.51	0.84	1.65	1.75	1.60	1.76

First, we compare the spectral power for each $y_{i_j}(t)$ and compare to the center position $y_{5_j}(t)$. The spectral power was calculated for frequencies in the range of [100, 1000] Hz, which is where the majority of the power lies in lung sounds [14]. An average power for each driver position i is calculated as $\bar{P}_i = \frac{1}{20}\sum_{j=1}^{20} P_{i_j}$, where the power of recorded signals P_{i_j} are averaged for all j trials. Each $\bar{P}_i$ is then compared to the average power from the center position $\bar{P}_5$, calculated as logarithmic ratio $\mathbf{P} = 10\log(\bar{P}_i/\bar{P}_5)$.

Second, an analysis of variance (ANOVA) test is done to determine whether there are any statistically significant differences between the full spectra of the driving and recorded signals at the nine positions for the three systems. To do this, the euclidean distance is measured between the magnitude spectrum of the recorded signal $|Y_{i_j}(n)|^2$ and the magnitude spectrum of the driving pink noise signal $|X_p(n)|^2$.

$$d_{i_j}(|Y_{i_j}|^2, |X_p|^2) = \sqrt{\sum_{n=1}^{N}(|Y_{i_j}(n)|^2 - |X_p(n)|^2)^2} \quad (2)$$

Considering the distance d_{i_j} for each of the i positions as a group, and each of the j trials for that position as observations for that group, we use a one-way ANOVA to determine if there a statistical significance in the distances between spectra based on position. In other words, this test determines whether the position of the input signal changes the output response of the systems.

B. Clinical Simulation

The evaluation of the quality of the body signals that is delivered to the end-user of the JHUscope is compared to other commercially-available stethoscopes in both quiet and extremely noisy environments. Noise robustness quantifies the deterioration in auscultation signal quality as a function of background noise. The proposed auscultation system was compared against six commercially available systems, including state-of-the art and widely used acoustic or electronic devices, as listed in Table II. These systems are chosen to represent a wide selection of options including acoustic and electronic devices as well as a wide range of features from static systems to active filtering ones. It is important to note that this selection is not meant as an exhaustive list of all stethoscopes on the market, nor is it a statement about the performance of these systems in clinical settings. It is rather a comparison of their individual performance under specific simulated conditions.

The noise robustness validation tests are performed in a simulated noisy clinical setting recreated inside a sound booth in order to control the acoustic environment. Real auscultation sounds from both normal and abnormal lungs are played from a

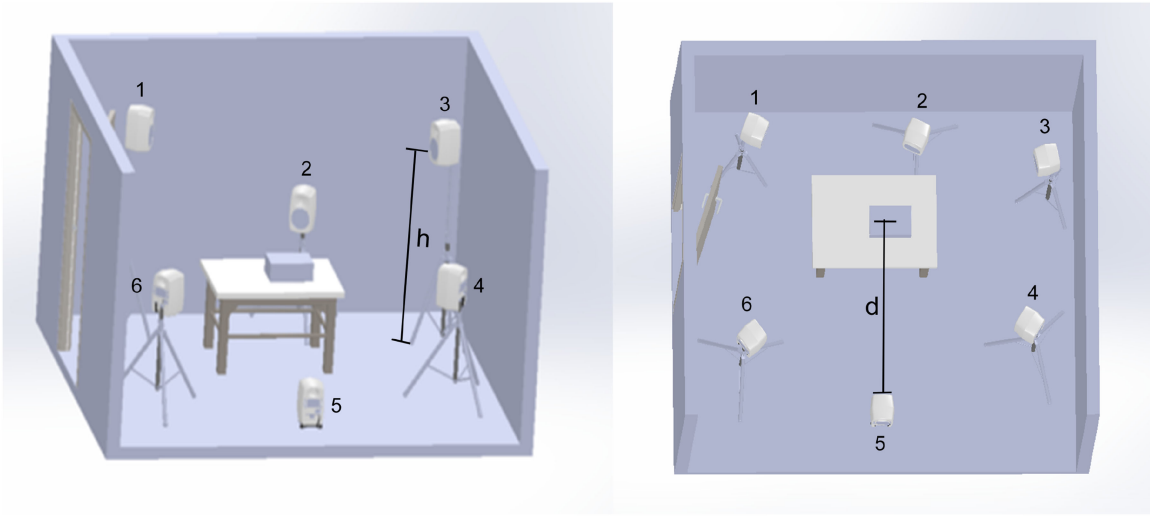

Fig. 3. Schematic of the experimental setup illustrating the placement of the loudspeakers and the chest sound simulator (rectangular prism) from the side (left) and the top (right). h is the loudspeaker's height position and d is the horizontal distance from the center of the chest sound simulator. The loudspeakers all face towards the chest sound simulator.

TABLE II
LIST OF STETHOSCOPE DEVICES AND SETTINGS

Device	Type	Filter	Vol./max
ADC Adscope	Acoustic	Diaphragm	N.A.
ADC Proscope SPU	Acoustic	Diaphragm	N.A.
EKO Core	Electronic	Diaphragm	3/ 7
Littmann Cardiology II	Acoustic	Diaphragm	N.A.
Littmann 3200	Electronic	Extended	3/ 7
Thinklabs One	Electronic	'Lung' Filter	5/ 10

chest sound simulation device, custom-built out of a loudspeaker covered in ballistic gelatin mimicking human tissue characteristics. The resulting signals emulate lung signals emanating from the body, used here as a reference for comparison. The use of this setup ensured: i) control over the volume level of the emanating body sounds and ambient noise, ii) control over the variety of the presented sounds iii) a repeatable setup for comparison among various simulated noise profiles, iv) uniformity in sound pick-up among auscultation systems, and v) reproducibility of results.

1) Clinical Room Simulation: Six loud speakers, one table, and one chest sound simulator were placed inside a sound booth of dimensions 148 in (long) x 123 in (wide) x 89 in (high), (Fig. 3). The table was positioned towards one end of the room while the six Genelec 6010 A loudspeakers were arranged to roughly face towards the table at different height placements (Table I), independently broadcasting noise sounds of various types and levels, reflecting a variety of examination settings from a quiet room to a busy clinic. On top of the table, the chest sound simulator transmitted low volume body signals. The noise and body sounds were delivered via a connected computer, stationed outside of the sound booth.

For a given noise background and condition (signal-to-noise level), a commercially available stethoscope or the JHUscope are positioned on top of the chest simulator in this setup and used to auscultate the reference signals while a calibrated microphone is placed directly above it to capture the ambient noise signal.

2) Chest Auscultation Simulation: A chest sound simulator transmitted digital breath signals at a low, fixed level (comparable to real chest auscultation signals) via a connected computer installed outside of the soundbooth. It was built comprising of a Jawbone Jambox loudspeaker with frequency response of 40–20 000 Hz and improved low-frequency sensitivity (via a proprietary bass radiator). The loudspeaker was covered in 1.5 in-thick medical synthetic gelatin (Humimic Medical, USA) that closely simulates the density and mechanical properties of human muscle tissue and can be kept at room temperature without deforming [20], [21]. The gelatin was baked at 200°F for 4 hours. A mold of the Jawbone Speaker was then placed inside the ballistic gelatin and it was allowed to cool at room temperature for 12 hours. The mold was replaced with the speaker and sealed with a heat gun; acoustic foam was placed at the bottom of the simulator to attenuate high frequency resonance and noise coming from areas of contact with the table.

Each auscultation system was placed individually at a single designated location on top of the chest simulator and held in position using a clamp to apply moderate pressure and ensure the setup remained the same throughout the completion of all simulations for each system.

3) Data Preparation: A collection of ten abnormal and ten control breath sounds of 10 s duration each were selected from a dataset of auscultation signals [22]. The abnormal group consisted of breath sounds containing wheeze, crackle and stridor sounds, while the control group consisted of mostly normal breath sounds recorded over various chest and tracheal areas. All digital clips were downsampled to 8 kHz.

The **High Stationary** group of ambient noises consisted of several colored noise subgroups including white, pink, violet, blue and brown, and fan-like noise found in the BBC database [23]. In total, 20 ten-second clips were selected from each subgroup, resulting in 120 High Stationary noise sounds.

The **Low Stationary** noise group consisted of noise types found in the BBC and NoiseX-92 databases [23], [24], and included subgroups of hospital ICU noise, hospital corridor noise, pulse monitor sounds, ambulance noise, babble noise and ambient talk, baby cry, street noise, chirping birds. Random silence periods were interjected to the noise clips to accentuate their non-stationary nature. An equal number of sounds were selected from each subgroup and sorted according to their average Power Spectral Entropy value (Eq. (3)), which was used as a stationarity index such that minimum entropy occurs for highly variable sounds (low stationary noises) and maximum entropy occurs when the spectral distribution in uniform (high stationary noises).

$$\overline{PSE} = \frac{1}{M}\sum_{m=1}^{M} PSE(m) \tag{3}$$

$$PSE(m) = -\sum_{k=1}^{K} p(m,k)\ \ln p(m,k) \tag{4}$$

$$p(m,k) = \frac{\widehat{PSD}(m,\omega_k)}{\sum_{k=1}^{K}\widehat{PSD}(m,\omega_k)} \tag{5}$$

where index $m = 1, \ldots, M$ signifies the processing windows of duration 100 ms, ω_k with $k = 1, \ldots, K$ and $K = 1024$ corresponds to the frequency index, and $\widehat{PSD}$ to the multitaper power spectral density estimate [25]. In a ranked list from lowest to highest entropy, the first 120 sounds were selected and attributed to the Low Stationary group; the rest were discarded.

4) Sound Playback: While the chest sound simulator emitted the selected body sounds at a fixed low volume, the loudspeakers independently broadcasted noise sounds randomly selected from the database. The volume of individual speakers was set to a random difference of $\{0, \pm 0.5, \pm 1, \pm 2\}$ dB from each other. The master speaker volume was automatically adjusted at the beginning of each trial, ensuring a net noise effect of signal-to-noise ratio at various levels, $SNR_{true} \in \{-20, -10, 0, 10, 15\}$ dB. All transmitted lung sound clips were pre-amplified independently to ensure equal average sound levels. In total 10 normal and 10 abnormal lung sounds were used, each assigned to 5 random net noise combinations (trials), resulting in a set of 100 sound recordings per SNR_{true} value: 50 abnormal and 50 normal.

The true SNR level, SNR_{true}, for each lung sound/net noise combination was determined by the short-term average of the ratio of the individual signal powers, averaged over M frames of duration 100 ms each:

$$\widehat{SNR}_{true} = \frac{1}{M}\sum_{m=1}^{M} 10\ \log_{10}\ \frac{P_s(n)}{P_d(n)} \tag{6}$$

where $P_*(n)$ is the average signal power of the n^{th} time frame; s corresponds to the sound signal recorded on the designated reference point on top of the chest simulator; d corresponds to the net ambient noise picked up adjacent to this designated point. Due to the dynamic nature of the lung sounds and noises as well as the randomly interjected silence periods, the calculation of SNR_{true} considered only the top 30% of frames in order to capture the highest average signal power. This way, it is ensured that no sound events exceed the desired SNR, while allowing for lower sound level events to be present.

5) Sound Capture: For the calculation of the true SNR in Section III-B4, signals of interest s and d were recorded independently (not simultaneously). Signals s and d were obtained using two 1/4" PCB Piezotronics prepolarized, omnidirectional condenser microphones, connected to a Brüel & Kjaer 5935-L preamplifier at 20 dB gain. The first microphone was placed on top of the chest simulator, facing downwards for recording signal s; and the second microphone was placed adjacent to the first with a separating distance of 10 cm, facing upwards, for recording signal d. The distance of these microphones from the speakers follows the positioning presented in Table I and it is assumed that the signal power from the speakers is equal at both microphones.

For the calculation of the quality metrics presented in Section III-C, simultaneous recordings were captured from the auscultation signal y, the reference sound driving the simulator x, and the ambient microphone d. Digital sounds from the auscultation system were captured using an 8-track ZOOM H4 recorder, situated outside the sound booth. All inherent sound effects and sound filters were disabled and the master recording gain was set at 0 dB.

C. Quality Metrics

The signal fidelity and noise robustness of each device is quantified by comparing the recorded auscultation signal y against the reference signal x that is driving the respiratory sound simulator as well as against the noise background d recorded by a calibrated microphone positioned above the chest sound simulator. This is done in order to assess two aspects of the signal: (1) how distorted the recorded signal is relative to the reference signal; and (2) how much leakage the device has in terms of letting noise mask the recorded auscultation.

There is currently no agreed upon standard for auscultation, so two existing objective quality metrics were chosen to quantify the similarity between the reference and measured signals and the amount of dissimilarity between the ambient and measured signals: normalized covariance measure (NCM) and magnitude squared coherence (MSC) [26]. Although NCM is a speech-based measure (as are most quality measures in the literature), it is biased by the perceptual system, accounting for signal audibility at various frequency bands for the human ear hence reflecting a general account of improved quality of a signal as perceived by a human listener, or in this case the auscultating clinician, on a scale of 0 (lowest quality) to 1 (highest quality). MSC is a statistical index that operates in the spectral domain

and examines the relation between two signals. Unlike NCM, this metric treats the full frequency spectrum evenly and linearly and the inclusion of MSC limits the shortcomings of using a perceptually biased metric alone. As with NCM, MSC gives a value of 1 for signals that have exact coherence to each other.

The metrics were chosen for their i) objective and standardized quality assessment ii) high correlation to human intelligibility scores, and iii) independence to signal amplification or volume variations of the signals. The measures also highlight two different aspects of the system: MSC values are not affected by non-linear operations whereas NCM is a measure that reliably predicts the effects of non-linear operations such as envelope thresholding or distortions introduced by spectral-subtraction algorithms [27].

1) Normalized Covariance Measure: NCM uses the covariance of the spectral envelopes of two signals x and y to calculate the Signal to Noise quantity SNR_{NCM}, weighted by band-importance factors W_j [26]

$$NCM(x,y) = 10\log_{10}\frac{\sum_{j=1}^{J} W_j\, \widehat{SNR}_{NCM}(x_j,y_j))}{\sum_{j=1}^{J} W_j} \quad (7)$$

The center frequencies for the 8 bands $j = 1,\ldots,J$ follow the Bark scale from 150 Hz to 4000 Hz. The corresponding bandwidths and band-importance weights W_j, follow the ANSI-1997 standards [28]. The Signal to Noise quantity SNR_{NCM} was calculated over M windows, as follows:

$$\widehat{SNR}_{NCM}(x_j,y_j) = \frac{(\sum_{m=1}^{M} V_x^{\mu}(m,j)V_y^{\mu}(m,j))^2}{1-(\sum_{m=1}^{M} V_x^{\mu}(m,j)^2 \sum_{m=1}^{M} V_y^{\mu}(m,j)^2)} \quad (8)$$

$$V_x^{\mu}(m,j) = V_x(m,j) - \mu_x(j) \quad (9)$$

where $V_x(m,j)$ is the centered spectral envelope of signal x for band j and window m, and $\mu_x(j)$ is the time average of envelopes $V_x(m,j)$ for all $m = 1,\ldots,M$ windows for signal x. The denominator comprizes of normalization factors that render quantity $\widehat{SNR}_{NCM}$ limited within [0, 1]; notice that NCM metric is also invariant to scalar multiplications of the input signals.

2) Magnitude Squared Coherence: MSC uses the normalized cross spectral energy to calculate a measure of coherence:

$$MSC(x,y) = 10\log_{10}\frac{1}{K}\sum_{k=1}^{K}\frac{|P_{xy}(\omega_k)|^2}{P_{xx}(\omega_k)P_{yy}(\omega_k)} \quad (10)$$

where $P_{xy}(\omega_k)$ is the cross-power spectrum density between signal x and y, with frequency spectrums $X(\omega_k)$ and $Y(\omega_k)$ respectively. For a given frequency band ω_k, the spectral density P_{xy} was estimated by $\hat{P}_{xy} = \sum_{m=1}^{M} X_m(\omega_k)Y_m^*(\omega_k)$ along $m = 1,..,M$ window frames, where M is determined by duration of window p, as described in Section III-C3. The denominator in (10) makes the MSC index normalized in [0,1]; notice that by definition, MSC is invariant to scalar multiplications of signals x and y; this renders the metric independent of the volume settings of the individual auscultation systems.

3) Overall Quality Metric: Both NCM and MSC are invariant to amplification of the input signals and obtain values between in the range of [0,1]: 0 when the signals under consideration have low similarity and 1 when the signals have high similarity. As a reference, a value of zero would be obtained if one of the compared signal originated from a white Gaussian process. NCM and MSC were computed over M non-overlapping Hamming windows, the durations of which varied from short to longer windows, $p = \{0.1, 0.5, 1, 2\}$ sec, to compare the measured output of the auscultation system y to the reference lung sound signal driving the chest simulator x and to the concurrent net ambient noise during auscultation d. An overall metric Q_p takes the geometric mean of the NCM and MSC metrics for each window length p and each x-y and d-y pairs:

$$Q_p(*,y) = \frac{NCM_p(*,y), +MSC_p(*,y)}{2} \quad (11)$$

Q_p is then averaged across all windows to create combined metrics accounting for the faithful representation of the emitted lung sounds ($SNR_{lungsound}$) and the amount of noise leakage into the auscultated signal (SNR_{noise}).

$$SNR_{lungsound} = \frac{1}{P}\sum_{p=1}^{P} Q_p(x,y) \quad (12)$$

$$SNR_{noise} = \frac{1}{P}\sum_{p=1}^{P} Q_p(d,y) \quad (13)$$

Finally, a single SNR metric, SNR_{est}, is calculated by subtraction of SNR_{noise} from $SNR_{lungsound}$. Since the signal metrics NCM and MSC (and therefore Q_p, $SNR_{lungsound}$, and SNR_{noise}) are calculated in the logarithmic scale, the subtraction of these metrics represents the division of linear-scaled $SNR_{lungsound}$ and SNR_{noise} metrics.

$$SNR_{est} = SNR_{lungsound} - SNR_{noise} \quad (14)$$

A graphical representation of the signal analysis methodology is included in Fig. 4.

D. Stethoscope Systems

Six commercially available acoustic and electronic stethoscope were chosen to represent construction and parameters of a majority of devices available currently; these parameters influence the characteristics of the transmitted sound and performance in noise settings. An overview of the systems is presented in Table II.

The ADC Adscope and Littmann Cardiology II devices were evaluated using the diaphragm chestpiece. The ADC Proscope single-patient use (SPU) device only entailed a diaphragm chestpiece. For these three devices, acquisition of the captured body sounds was performed using the PCB Piezotronics microphone, secured into one side of the earpiece. Both sides of the earpiece were then acoustically sealed via a 3-step process: i) covered with a thick layer of clay ii) wrapped in multiple layers of acoustic foam, iii) wrapped in cotton cloth. This process ensured

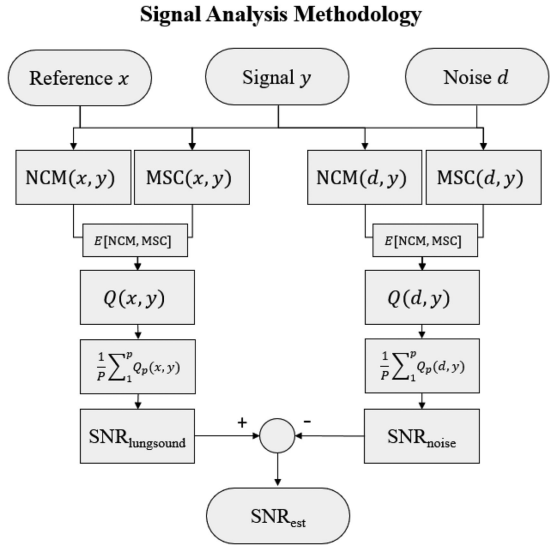

Fig. 4. Flowchart highlighting the signal analysis methodology for determining the overall quality metric SNR_{est}.

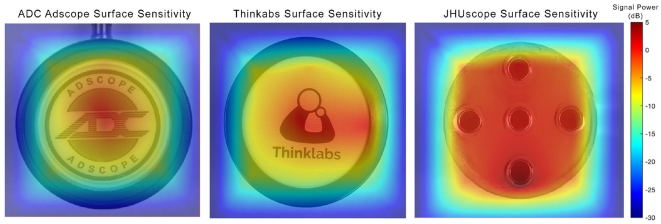

Fig. 5. Sensitivity map of a traditional stethoscope head (left), sensitivity map of the Thinklabs One electronic stethoscope head (center), and sensitivity map of the proposed JHUscope smart stethoscope (right) in decibels as compared to the power at the center position.

restriction of potential noise leakage through the ends of the earpiece. It is important to highlight here that such systems for which sound travels through a chestpiece, noise leakage is more than likely to occur throughout the full length of the tubing piece. It was outside of the scope of this study to attempt to contain all possible sources of noise leakage, especially those that add to a device's vulnerabilities.

The EKO Core hybrid electronic/acoustic device was toggled to *ON*, to ensure digital acquisition, and the middle volume setting was selected while the diaphragm chestpiece was used for auscultation. Sound was acquired through the earpiece using the same process as above. Notice here that the device offers an accompanying phone application for digital sound capturing and while the Bluetooth indicator was flashing, the device was not connected to a phone to ensure uniformity in the recording process. The current study focused on assessing the audio signal reaching the user in real-time, simulating a scenario of real-world auscultation; and thus, additional computer software was not considered here.

The Littmann 3200 electronic device was set to active mode (non-standby mode), the filter option was set to Extended mode, volume set at the middle setting. The Extended Range mode amplifies sounds from 20–2000 Hz similar to the Diaphragm Mode, but provides more low frequency response between 50–500 Hz. Sound was again acquired using the process above. Due to the restricted automatic shut-off feature, the device had to be set into active mode regularly throughout the duration of the experiments. This device also comes with an accompanying computer software for digital sound acquisition which was also forgone to ensure uniformity in the recording process and a more realistic use-case.

Thinklabs One electronic device offers a standard 3.5 mm audio jack output where the user connects headphones, providing the capability of directly recording the transmitted sound via an audio cable connected to the recording system. The filter option recommended for lung sounds was used (filter setting 3-4), and the volume was set at the middle point.

IV. Results

A. Sensing Sensitivity

First, we measure the output power as a function of location. Fig. 5 shows a surface sensitivity heatmap of the JHUscope versus the ADC acoustic stethoscope and the Thinklabs One electronic stethoscope, revealing spectral power in decibels relative to the center position. In a traditional stethoscope head, the diaphragm is designed to partially integrate over the surface of the head [29]. However, our analysis shows that the acoustic stethoscope is maximally sensitive at the center of the stethoscope head, directly under the opening leading to the tubing, and decreases dramatically towards the edges with a loss of almost 30 dB in dynamic range at the outer edges of the stethoscope (Fig. 5-left). Similarly, the Thinklabs One electronic stethoscope exhibits positional relationship to signal power (Fig. 5-center), possibly due to transducer design. In contrast, the JHUscope increases the number of pickup positions across the stethoscope's head with a five microphone array and therefore provides a more uniform surface to capture body sounds (Fig. 5-right). These results are consistent with previously reported studies that demonstrate that diaphragms have the ability to increase the mean vibrational velocity across the surface but have significant differences between the velocity values measured at center and edge points of those diaphragms [30].

In order to further illustrate these differences and the power loss at certain positions, Fig. 6 shows an example of the power spectrum from the JHUscope and the acoustic systems for a position that is distant from the center. In this example, the average power spectrum from Position 1 relative to the average power spectrum from the center position, Position 5, was chosen to illustrate the spatial dependencies of the sensitivity. For this position, the JHUscope preserves signal power relative to the center and across the frequencies of interest while the ADC Adscope shows a significant decrease in power—down to 20 dB less—and high variability across the spectrum.

Second, we further evaluate whether the full spectra of the output and input signals are statistically different ($p < 0.05$) across pickup positions. An ANOVA parametric analysis reveals

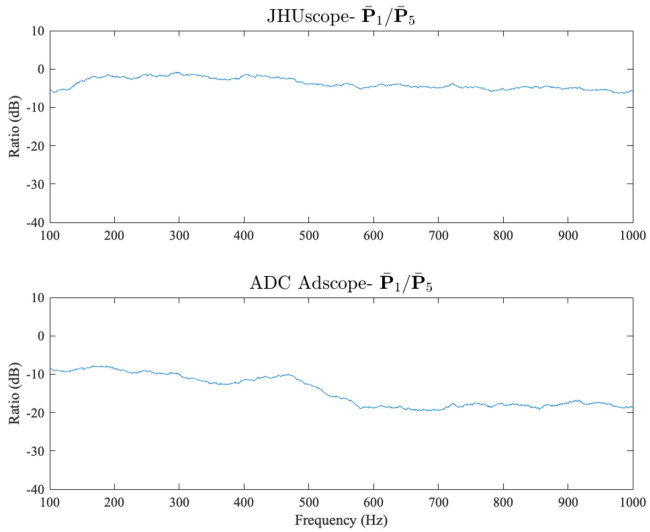

Fig. 6. Example of spectral power differences between the JHUscope and ADC Adscope, illustrated by the spectral power from Position 9 relative to the center position, Position 5. The spectral power ratio is presented in decibels (dB), with 0 dB representing equal power of recordings from both positions.

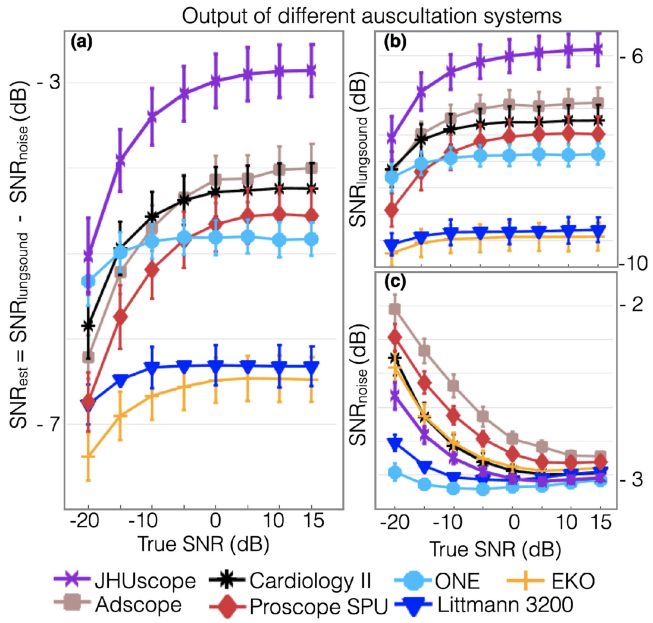

Fig. 7. Illustration of the sound-preservation ability of different auscultation systems, with varying simulated noise levels. The true SNR is depicted on the x-axis and the estimated SNR is plotted on the y-axis. Main panel (a) depicts results for calculated metric SNR_{est}; panel (b) depicts the $SNR_{lungsound}$ metric, and panel (c) the SNR_{noise} metric. In panels (a–b) high values correspond to high quality of the pick-up signal; in panel (c) high values correspond to maximal noise leakage.

that the recorded and true signal for the JHUscope are found to be statistically equivalent across positions ($p = 0.1204$, $F = 1.67$) but are statistically different for the both the ADC Adscope ($p = 0.0242$, $F = 7.14$)and the Thinklabs One ($p = 3.8053E-27$, $F = 27.53$) across positions.

Several publications have studied the distribution of sounds of the chest wall, demonstrating the presence of sounds across large areas, with minor amplitude differences ($\pm$ 3 dB) based on the placement of the sensor [31], [32]. One of the key advantages of this sensing array design is that the JHUscope does not require precise placement on the body to achieve the maximum amplitude, as it is uniformly sensitive across the head of the device, thereby allowing for increased sound and information pickup for effective use by health-care workers with minimal training. In the cases with highly trained personnel requiring more localized sensitivity, the programmable aspect of the JHUscope enables selective toggling of the microphones, the comparison of which is shown in Fig. 1 of the Supplementary Material.

B. Signal Fidelity and Noise Robustness

Fig. 7(a) depicts estimated SNR values for each of the tested stethoscopes as a function of the true SNR (averaged across all trials, lung sounds and noise signals). High values of the SNR_{est} metric represent increased signal fidelity with the reference signal and reduced noise leakage, while low values reflect decreased signal quality and increased noise contamination. The overall metric SNR_{est} shown in Fig. 7(a) shows that the JHUscope far surpasses the other devices for all simulated conditions. Figs. 7(b), (c) show the breakdown achieved by the $SNR_{lungsound}$ and SNR_{noise} metrics alone, respectively.

As depicted in Fig. 7, it is clear that some devices perform well in quiet conditions, while others are built to withstand noisy environments where it may be more important to suppress ambient noise leakage than preserving the real signature of the breath sounds (e.g., electronic device Thinklabs One, in cyan). Moreover, it is clear that acoustic stethoscopes perform well in low noise conditions but cannot provide sufficient noise suppression capabilities in noisy settings (Fig. 7(c)). In contrast, electronic stethoscopes provide advanced filtering to suppress ambient noise (low SNR_{noise} values) but such filtering can also affect the underlying signature of the reference breath sounds (Thinklabs One in cyan, Littmann 3200 in blue and EKO Core in yellow). The proposed JHUscope device shows a balance between noise suppression of the background and signal fidelity with the reference sound for an overall improved SNR_{est}.

In order to further illustrate this point, Fig. 8 shows the spectrogams of a reference signal containing an abnormal lung sound (wheezing) and a simulated nonstationary ambient noise (ambulance noise). These spectrograms are compared to the spectrograms of recordings from four devices (JHUscope, ADC Adscope, Thinklabs One, and Eko Core) at three noise levels (SNR -10 dB, 0 dB, and 10 dB). The spectrograms clearly illustrate that some devices are better at signal preservation while other devices are better at noise cancellation. The ADC Adscope shows a better preservation of the wheezing across all three noise levels (high $SNR_{lungsound}$), but also a high level of noise contamination from the ambulance sounds (high SNR_{noise}). The Thinklabs One shows a very high level of noise suppression of the ambulance sounds (low SNR_{noise}), but at the expense of a substantial amount of the frequency content of the lung sounds, removing the high pitch wheezes entirely (low $SNR_{lungsound}$). The Eko Core shows a a high

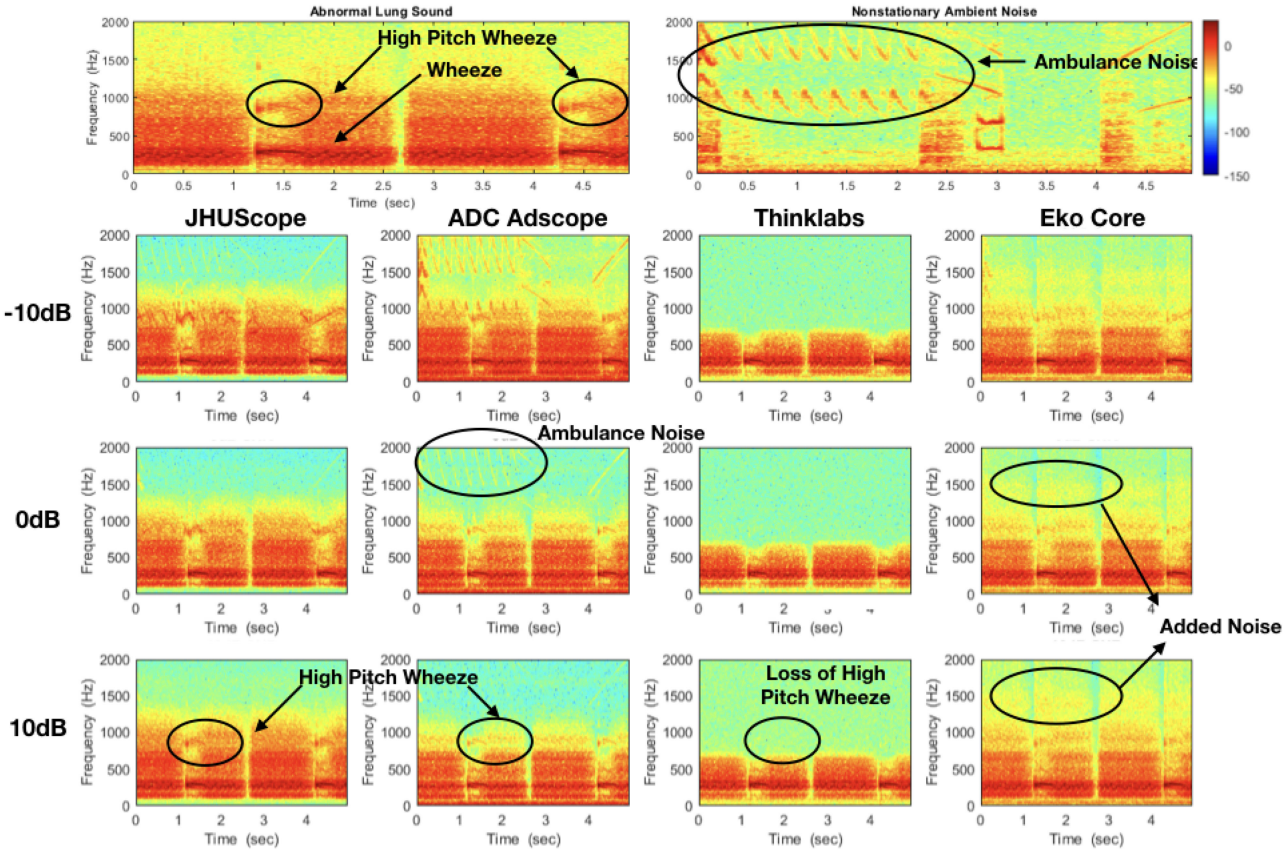

Fig. 8. Example of recordings from four different auscultation systems (JHUscope, ADC Adscope, Thinklabs One, and Eko Core) at three simulated noise levels, to highlight the issues summarized in Fig. 7. Panel (a) represents the spectrogam of the reference recording of breaths with expiratory wheezing, an abnormal lung sound signal. Panel (b) represents the simulated ambient noise spectrogram of nonstationary, transient sounds (ambulance noises). Panels (c)–(e) show the pick-up signal for the three devices for SNRs -10 dB, 0db, and 10 dB, respectively. The recordings from each device can be compared to the reference spectrograms to illustrate the $SNR_{lungsound}$ and SNR_{noise} metrics.

level of noise suppression (low SNR_{noise}), but adds noise in higher frequency bands, resulting in a low $SNR_{lungsound}$. The proposed JHUscope preserves the power of the abnormal lung sounds in all three scenarios (high $SNR_{lungsounds}$) while also being able to suppress the ambulance sounds (low SNR_{noise}).

V. Discussion

Three main points should be taken into consideration with the current study. First, we note that there is expected to be variability across system performances that is driven by a variety of factors including product design and the physical device, system architecture and inherent signal processing, and signal acquisition layout. The quality performance results presented in this work are based on the standalone capabilities of the included devices, and do not include supplementary computer programs or phone applications in the analysis.

Some well-known sources of sound alteration are the chest piece design and the attached tubing typically found in acoustic stethoscopes and some electronic devices (e.g., ADC Adcope, Littmann Cardiology II, EKO Core, Littmann 3200), as well as digital filters in electronic systems. The diaphragm of acoustic and electronic stethoscopes, by design, is meant to cut off some of the very low and very high frequencies [33]. The narrow tubing and earpiece components are another source of sound alteration [34], [35]: sounds propagating through narrow long tubes have specific resonant frequencies and a decreased response to high frequency content. In addition to sound alteration, the chestpiece and tubing may further incur ambient noise leakage, especially in noisy environments. This can result in a contaminated propagating sound that drifts further away from the actual body sounds. Electronic system also have various sources of sound alteration. Electronic stethoscopes typically utilize various types of electroacoustic contact transducers that will influence the sensitivity and sound characteristic of the pickup signal in a different manner than a diaphragm of an acoustic stethoscope [36]. Digital filters in electronic systems that are geared towards lung sound auscultation typically suppress frequencies below 300 Hz and above 800–1000 Hz, and although these are deliberate design choices for attenuating external noise and promoting specific signal frequencies, they may also further contribute to a decreased signal fidelity in certain electronic stethoscopes.

Second, the performance curves should not be interpreted as an absolute ranking of the various auscultation systems. Instead, they provide an insight into the variability of each individual

system when the level and type of the ambient noise varies. Note that a system's output can be sensitive to the amount of pressure used when placed on the body or on top of the chest simulator. High pressure allows for better contact with the chest simulator and less ambient noise leakage via the chest piece. In our setup, the use of clamps to secure the chest piece of each device ensured minimal setup variability; however, due to the different shapes and sizes of the auscultation devices full elimination of pressure variability could not be guaranteed in a straightforward manner. In addition, there is a great deal of variability in setting choices including volume of each system. Although steps were taken to prevent such variability such as using the middle volume setting for all electronic devices and choosing performance metrics to be invariable to scalar multiplications of the signals, one could argue that the volume setting might still affect the end result: a high volume setting might allow for less relative noise leakage that can contaminate the pick-up signal or it could be the case that some systems process incoming sounds in a different way dependent on the input volume level.

Lastly, the objective metrics in this study reflect the *systems'* ability to preserve the reference emanating body sounds, but should not be used to reflect an expert's ability to form a diagnostic opinion using these systems without further exploration. The present study compares the auscultated signals against the true breath sounds driving the chest simulator through a metric that quantifies the fidelity of the information contained in the auscultation signal rather than other metrics such as amplification. Indirect measures that focus on sound quality rather than fidelity can obscure any filtering, attenuation, or distortions applied by the stethoscope to abnormal body sound events. A direct correlation to the end-user's ability to diagnose a body condition based on the delivered sounds would require further evaluation with expert listener panels and would depend on the nature of the body sounds of interest.

In future investigations, the real-time noise suppression algorithm should be validated against other noise reduction methodologies in a similar manner to further characterize and understand the limitations of the current methodology, such as the use of the Short-time Fourier Transform. We should also do a full comparison of the cost, processing power, battery life, and programmability against other available devices. To fully validate the signal quality of the device, we would also need to consider real-life conditions. Although a preliminary study has been conducted of a pairwise comparison of the JHUscope and the Thinklabs One in a small pediatric population [37], the JHUscope needs to be tested on a more extensive sets of normal and abnormal lung sounds, a larger patient population with various conditions, and in various real-life environmental noise conditions. The system's validation would also benefit from evaluations for use with other body sounds, such as heart, bowel, and joint sounds.

VI. Conclusion

This work describes a new digital stethoscope that brings solutions to old and recurring problems of auscultation tools by equipping the system with an advanced sensing mechanism, dynamic noise suppression design, and programmability to tackle known shortcomings. Noise leakage and sound alteration effects are evident among all compared systems. Although high noise suppression is generally desirable, it can incur significant sound alteration to the auscultated sounds. A balanced solution is thus desirable which maintains the full spectrum of the sound while minimizing noise. Our experiments have shown that the JHUscope delivers sound signals faithful to the reference body sounds, achieving increased pick-up sensitivity and decreased noise leakage. The proposed system is proven to be a robust, powerful and versatile tool that can be used reliably in challenging and noisy environments. This system can bring value to the clinicians in their traditional auscultation process, but can also add value to computer-aided auscultation systems (CAAS) that would benefit from a broader and more accurate representation of body sounds. Most CAAS approaches have mainly been validated in well-controlled or quiet clinical settings on adult subjects; there is yet to have a true impact on health-care practices in various clinical settings [38]–[41].

Since lung auscultation remains an important component of respiratory infection diagnosis with more predictive accuracy than an initial clinical assessment alone [5], a device that can provide clinical workers with markedly improved signal quality may be critical for increasing the accuracy of clinical diagnosis of respiratory infections. Better still, the programmability of the JHUscope can enable deployment of real-time CAAS for clinical decision support: previously-evaluated CAAS algorithms that can subsequently be introduced on this hardware device [42], [43] would be able to be deployed to both traditional and non-traditional clinical settings and support the diagnosis of respiratory conditions, bringing patients closer to successful treatments in resource-limited settings.

Competing Interests: The authors declare the existence of a financial competing interest. Under a license agreement between Sonavi Labs and the Johns Hopkins University, I.M. and the University are entitled to royalty distributions related to technology used in the study discussed in this publication. Additionally, the University owns equity in Sonavi Labs. I.M. is a founder of and holds equity in Sonavi Labs and serves as the company's Chief Technology Officer and a member of its Board of Directors. This arrangement has been reviewed and approved by Johns Hopkins University in accordance with its conflict of interest policies.

References

[1] T. Ferkol and D. Schraufnagel, "The global burden of respiratory disease," *Ann. Amer. Thoracic Soc.*, vol. 11, no. 3, pp. 404–406, 2014.
[2] A. Sakula, "R t h laënnec 1781–1826 his life and work: A bicentenary appreciation." *Thorax*, vol. 36, no. 2, pp. 81–90, 1981.
[3] J. S. Russotti, R. P. Jackman, T. P. Santoro, and D. D. White, "Noise reduction stethoscope for united states navy application," *Nav. Submarine Res. Lab, Groton, CT*, Tech. Rep. 1214, 2000.
[4] World Health Organization, *Revised WHO classification and treatment of childhood pneumonia at health facilities*. Geneva, Switzerland, 2014.
[5] F. Pervaiz *et al.*, "Building a prediction model for radiographically confirmed pneumonia in peruvian children: From symptoms to imaging," *Chest*, vol. 154, no. 6, pp. 1385–1394, 2018.
[6] G. Nelson, "Stethoscope design for auscultation in high noise environments," Ph.D. dissertation, Univ. Minnesota, Minneapolis, MN, USA, 2015.

[7] J. Solà *et al.*, "Towards an unsupervised device for the diagnosis of childhood pneumonia in low resource settings: Automatic segmentation of respiratory sounds," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Oct. 2016, pp. 283–286.

[8] A. Poreva, Y. Karplyuk, A. Makarenkova, and A. Makarenkov, "Application of bispectrum analysis to lung sounds in patients with the chronic obstructive lung disease," in *Proc. IEEE 34th Int. Sci. Conf. Electron. Nanotechnol.*, 2014, pp. 306–309.

[9] M. Lozano, J. A. Fiz, and R. Jané, "Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 2, pp. 486–497, Mar. 2016.

[10] G. Nelson and R. Rajamani, "Accelerometer-based acoustic control: Enabling auscultation on a black hawk helicopter," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 994–1003, Apr. 2017.

[11] P. C. Loizou, *Speech Enhancement: Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2013.

[12] G. Prasad *et al.*, "A review of different approaches of spectral subtraction algorithms for speech enhancement," *Curr. Res. Eng*, vol. 1, no. 2, pp. 57–64, 2013.

[13] D. Emmanouilidou, E. D. McCollum, D. E. Park, and M. Elhilali, "Adaptive noise suppression of pediatric lung auscultations with real applications to noisy clinical settings in developing countries," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 9, pp. 2279–2288, Sep. 2015.

[14] L. E. Ellington *et al.*, "Developing a reference of normal lung sounds in healthy peruvian children," *Lung*, vol. 192, no. 5, pp. 765–773, 2014.

[15] O. S. Levine *et al.*, "The pneumonia etiology research for child health project: A 21st century childhood pneumonia etiology study," *Clin. Infect. Dis.: Official Pub. Infect. Dis. Soc. Amer.*, vol. 54, no. Suppl 2, pp. S 93–S101, Apr. 2012.

[16] S. B. Patel *et al.*, "An adaptive noise reduction stethoscope for auscultation in high noise environments," *J. Acoust. Soc. Amer.*, vol. 103, no. 5, pp. 2483–2491, 1998.

[17] G. Nelson, R. Rajamani, and A. Erdman, "Noise control challenges for auscultation on medical evacuation helicopters," *Appl. Acoust.*, vol. 80, pp. 68–78, 2014.

[18] N. Lu, C. Lu, S. Yang, and J. Rogers, "Highly sensitive skin-mountable strain gauges based entirely on elastomers," *Adv. Funct. Mater.*, vol. 22, no. 19, pp. 4044–4050, 2012.

[19] L. J. Nowak and K. M. Nowak, "Acoustic characterization of stethoscopes using auscultation sounds as test signals," *J. Acoust. Soc. Amer.*, vol. 141, no. 3, pp. 1940–1946, 2017.

[20] C. Van Sligtenhorst, D. S. Cronin, and G. Wayne Brodland, "High strain rate compressive properties of bovine muscle tissue determined using a split hopkinson bar apparatus," *J. Biomech.*, vol. 39, no. 10, pp. 1852–1858, 2006.

[21] M. L. Fackler and J. A. Malinowski, "Ordnance gelatin for ballistic studies. detrimental effect of excess heat used in gelatin preparation," *Amer. J. Forensic Med. Pathol.*, vol. 9, no. 3, pp. 218–219, Sep. 1988.

[22] J. A. S. McCann, *Nursing Know-How: Evaluating Heart & Breath Sounds.* Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2009, iD: 731183533.

[23] BBC, "Sound Effects," British Broadcasting Corporation. Accessed: Jan. 17, 2017. [Online]. Available: http://sound-effects.bbcrewind.co.uk

[24] A. P. Varga, H. J. M. Steeneken, M. Tomlinson, and D. Jones, "The noisex-92 study on the effect of additive noise on automatic speech recognition," Tech. Rep., Speech Res. Unit, Defense Res. Agency, Malvern, U.K., 1992.

[25] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol. - Heart Circul. Physiol.*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[26] J. Ma, Y. Hu, and P. C. Loizou, "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions." *J. Acoust. Soc. Amer.*, vol. 125, no. 5, pp. 3387–405, May 2009.

[27] R. L. Goldsworthy and J. E. Greenberg, "Analysis of speech-based speech transmission index methods with implications for nonlinear operations," *J. Acoust. Soc. Amer.*, vol. 116, no. 6, pp. 3679–3689, 2004.

[28] American National Standards Institute. *American National Standard: Methods for Calculation of the Speech Intelligibility Index*. Acoustical Society of America, 1997.

[29] T. Rossinger, *Handbook of Acoustics*, 2nd ed., T. Rossinger, Ed. Berlin, Germany: Springer, 2014.

[30] L. J. Nowak and K. M. Nowak, "An experimental study on the role and function of the diaphragm in modern acoustic stethoscopes," *Appl. Acoust.*, vol. 155, pp. 24–31, 2019.

[31] G. Benedetto, F. Dalmasso, and R. Spagnolo, "Surface distribution of crackling sounds," *IEEE Trans. Biomed. Eng.*, vol. 35, no. 5, pp. 406–412, May 1988.

[32] Y. Kawamura, Y. Yokota, and F. Nogata, "Propagation route estimation of heart sound through simultaneous multi-site recording on the chest wall," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 2875–2878.

[33] P. Y. Ertel, M. Lawrence, R. K. Brown, and A. M. Stern, "Stethoscope acoustics," *Circulation*, vol. 34, no. 5, pp. 899–909, 1966.

[34] G. Kirchhoff, "Ueber den einfluss der wärmeleitung in einem gase auf die schallbewegung," *Annalen Der Physik*, vol. 210, no. 6, pp. 177–193, 1868.

[35] D. E. Weston, "The theory of the propagation of plane sound waves in tubes," *Proc. Phys. Soc. B*, vol. 66, pp. 695–709, Aug. 1953.

[36] S. Leng, R. San Tan, K. T. C. Chai, C. Wang, D. Ghista, and L. Zhong, "The electronic stethoscope," *Biomed. Eng. Online*, vol. 14, no. 1, pp. 1–37, 2015.

[37] S. Graceffo, A. Hussain, S. Ahmed, E. D. McCollum, and M. Elhilali, "Validation of auscultation technologies using objective and clinical comparisons," in *Proc. Annu. Conf. Eng. Med. Biol. Soc.*, 2020, pp. 992–997.

[38] N. Q. Al-Naggar, "A new method of lung sounds filtering using modulated least mean square-adaptive noise cancellation," *J. Biomed. Sci. Eng.*, vol. 6, pp. 869–876, 2013.

[39] K. K. Guntupalli, P. M. Alapat, V. D. Bandi, and I. Kushnir, "Validation of automatic wheeze detection in patients with obstructed airways and in healthy subjects," *J. Asthma*, vol. 45, no. 10, pp. 903–907, Jan. 2008.

[40] J. Li and Y. Hong, "Wheeze detection algorithm based on spectrogram analysis," *Proc. 8th Int. Symp. Comput. Intell. Des.*, vol. 1, 2015, pp. 318–322.

[41] M. Yamashita, M. Himeshima, and S. Matsunaga, "Robust classification between normal and abnormal lung sounds using adventitious-sound and heart-sound models," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2014, pp. 4418–4422.

[42] D. Emmanouilidou, K. Patil, J. West, and M. Elhilali, "A multiresolution analysis for detection of abnormal lung sounds," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2012, pp. 3139–3142.

[43] D. Emmanouilidou, E. D. McCollum, D. E. Park, and M. Elhilali, "Computerized lung sound screening for pediatric auscultation in noisy field environments," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 7, pp. 1564–1574, Jul. 2018.