31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

On Arithmetic Misconceptions of Spectral Analysis of Biological Signals, in Particular Respiratory Sounds

Azadeh Yadollahi, *Student Member, IEEE* and Zahra Moussavi, *Senior Member, IEEE*

***Abstract*— Spectral analysis is one of the most common methods in sound signal analysis for approximating sound power. However, since the sound power is usually presented in logarithmic scale, it is important to consider the non-linearity effects of logarithm function. In this study, the misconceptions and implementation issues regarding noise power reduction and average power calculation are described. Respiratory sound analysis is utilized as an example to show these issues in a practical application. The results indicate that most of the errors happen during noise power reduction; they can be either due to substituting noise reduction by sound detection concept or/and representing the noise power in the very low frequency components instead of the signal power. Also, if the average powers of the signals are calculated in the wrong scale, the results do not represent the acoustical characteristics of the sounds; this is shown by considering the flow-sound relationship at different flow rates.**

I. INTRODUCTION

Spectral analysis of sound signals is one of the most common practical means to obtain useful information regarding the pressure, intensity and power of the sound signals at different frequency ranges. Respiratory sound signal is an example of biological signals for which the spectral analysis is frequently used to extract information, i.e. average power at different frequency ranges of the sound signals.

Since the intensity and power of sound signals change in a wide range of values, usually they are being presented in log scale to narrow the range. However, due to the non-linear characteristics of log function, special issues should be considered when it is applied along with other arithmetical functions on the power values of the sound signals. Otherwise, the results may not represent the acoustical properties of the sound signals.

In this paper we show the details of mathematical implementations when combining log and other arithmetical functions for respiratory sounds analysis purpose. To show the effects of different implementations of these functions, two applications of noise power removal and respiratory flow-sound relationships are examined and compared.

A. Basic Definitions

When recording sound with a microphone, the output of the microphone (in volts) shows the changes in the sound pressure at the surface of the microphone, which is related to the power (P) of the sound [1]. Spectral analysis of the recorded sound signals is used to estimate the power components of the sounds at different frequency ranges. Most commonly, Fast Fourier Transform (FFT) is applied to the signals to estimate either power spectral density (PSD) of the signals in terms of Watts, and then is shown in the log scale in dB.

For non–stationary signals, i.e. respiratory sounds or speech signals, the spectrogram of the signal is calculated by short time Fourier Transform (STFT) in short-duration overlapping windows and the average power of the signal in dB is defined as $AvgPwr = 10\log\left(E[P_i]/P_{ref}\right)$, where $E[P_i]$ shows the average of signal's power. P_{ref} is a reference power such as $10^{-12}W$, i.e. the smallest audible power of a pure 1kHz tone. The index i refers to segments in which their average power is of interest.

If P_s and P_n show the power of sound and noise (in Watts), respectively then, the minimum ratio of P_s/P_n that guarantees detection of sound in the receiver (ear) is specified by the detection threshold, $DT = 10\log\left(P_s/P_n\right)$, [1], which is known as the signal to noise ratio, SNR. This is a commonly used definition, which is particularly important for investigating the audibility of sound signals in the presence of acoustic noise.

II. RESPIRATORY SOUNDS ANALYSIS

Figure 1 shows the schematic of the steps performed in most of the studies in which time-frequency analysis of respiratory sounds is used. Respiratory sounds are normally band-pass filtered to remove very low (below 50 Hz) and high frequency components of noise. FFT is then applied to the filtered signals to estimate either PSD or STFT of the signals in terms of Watts or dB.

A. Noise removal

An important step in pre–processing of the respiratory sounds is noise removal. When recording respiratory sounds, ambient noise is also recorded, which may contain components within the frequency range of respiratory sounds. Assuming that the noise is additive, it can be easily shown that the power of the recorded signal in each frequency range (i) equals to $P_i = P_{Si} + P_{Ni}$, where P_i, P_{Si} and P_{Ni} are powers of the recorded signal, respiratory sounds and the ambient noise in the same frequency range (i), respectively.

To remove the ambient noise components from the PSD or STFT of the recorded signal, it is required to have an approximation of the noise power at different frequency

A. Yadollahi is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada, R3T 5V6. She is also a research assistant in TRLabs Winnipeg, 100-135 Innovation Drive, Winnipeg, MB, Canada. azadeh@ee.umanitoba.ca

Z. Moussavi is with the Department Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada, R3T 5V6. She is also an adjunct scientist in TRLabs Winnipeg, 100-135 Innovation Drive, Winnipeg, MB, Canada. mousavi@ee.umanitoba.ca

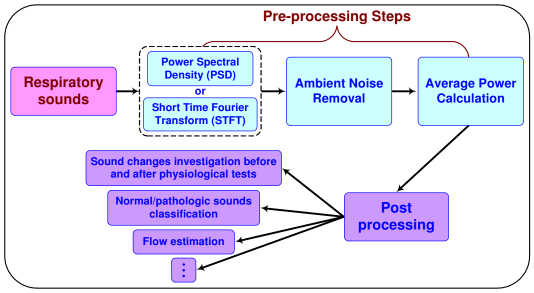


Fig. 1. Schematic of respiratory sounds analysis based on average power calculation.

ranges. Therefore, usually at the end of each recording of respiratory sounds the subject is asked to stop breathing for a few seconds. The PSD of the recorded signal during the breath hold period gives an estimation of the ambient noise power components at different frequency ranges.

The estimated power of the respiratory sounds in the log scale (dB) is:

$$dBP_{Si} = 10 \log (P_i - P_{Ni}), \quad (1)$$

When using the log transform of PSD and STFT values, there are a few issues to be considered in order to have a proper approximation of the respiratory sounds power that are explained below.

1) Subtracting in the right scale: Instead of subtracting the power values of noise in Watts from those of the recorded signal, if their power values in dB are subtracted, it would be equal to dividing the power values of the recorded signal by those of the noise:

$$dBP_i - dBP_{Ni} = 10 \log P_i - 10 \log P_{Ni} = 10 \log P_i/P_{Ni}, \quad (2)$$

where dBP_i and dBP_{Ni} are the powers of the recorded signal and noise in dB, respectively. It is clear that Eq. (2) does not yield the proper approximation of the respiratory sound power presented in Eq. (1).

2) Sound detection versus noise reduction: When removing the components of noise power, another error arises from the misinterpretation of sound detection threshold, or SNR in dB, and the noise removal concept. The former describes the smallest threshold for the sound to be audible in the presence of noise, while the latter is dealing with removing the noise power components from those of the recorded signal. Following the same reasoning discussed in previous section, it is evident that the sound detection threshold measure (SNR in dB) cannot be applied to remove the components of noise power from those of the recorded signal. However, SNR values can be used to find the segments where the respiratory sounds are audible in the presence of noise.

3) Mathematical implementations: Logarithm is an incremental function, which yields real values for positive real numbers in the range of $(0, \infty)$. Some computational softwares such as MATLAB calculate complex logarithm and return complex values for either negative or complex inputs. In the case of negative inputs, the output will be a complex number where:

$$Real\{\log(a)\} = \log(-a), \quad (3)$$

where $a < 0$, and $Real\{.\}$ returns the real part of a complex number. However, when displaying the results, MATLAB ignores the imaginary parts of the complex numbers and shows only the real values. Therefore, for the cases where the noise power values are greater than those of the recorded signals, such as low frequency components, special implementation details should be considered. Otherwise, instead of the power differences of $P_i - P_{Ni}$, the power differences of $P_{Ni} - P_i$ will be displayed.

B. Average power calculation

In many applications, researchers are interested to investigate the average power of the respiratory sound signal over different frequency ranges of interest. To calculate the average power, power components should be averaged in Watts, then transformed to the log scale and displayed in dB. However, it is a common mistake to average the values of signals power in dB over a frequency range, i.e. $[f_1, f_2]$. It can be easily shown that in this case the output would be the logarithm of the geometric average of the power values over the frequency range of $[f_1, f_2]$. Although, the geometrical averaging of the power components may have some signal processing applications, it does not convey a reasonable physical interpretation.

To compare the effects of these averaging schemes, $AvgPwr$ (average of power in Watts) and $AvgLog$ (average of power in dB) values of tracheal sound signals are used to describe the relationship between tracheal sound and flow at different flow rates in the following sections.

C. Data

Respiratory signals of 4 healthy adults were adopted from a previous study [2]. The recorded respiratory sounds were band-pass filtered in the range of $[50 - 2500]$ Hz to remove low and very high frequency components of noise. The subjects were instructed to breathe at low, medium, high and very high flow rates. At the end of each recording the subject stopped breathing for $10s$ to record an estimation of the ambient noise.

D. Flow-Sound relationship

Tracheal sound signal of each subject was normalized to have zero average and unity energy. Spectrum of the signal was calculated in windows of 100 ms (1024 samples) with 90% overlap between adjacent segments to have a smooth approximation of STFT. The power components of tracheal sound in the frequency range of $[200 - 800]$ Hz were considered as this range is free of the main components of heart sounds while including the main components of the tracheal sound [3].

Tracheal sound signals are non–stationary in nature [3]. To overcome this problem, in each respiratory cycle, only the segments corresponding to the upper (lower) 20% of the target flow of inspiratory (expiratory) phases are investigated, for which the respiratory sounds can be considered stationary in wide sense within those periods. On the other hand, the mechanisms of sound generation and the flow-sound relationship during inspiratory and expiratory phases are different [4], [5]; hence, each phase was investigated separately.

For each subject, the recorded airflow signal was used to distinguish respiratory phases and also marking the respiratory sounds corresponding to each of the low, medium, high and very high flow rates. For the respiratory cycles with similar flow rates, the relationship between the AvgPwr and AvgLog values of tracheal sound and the corresponding flow signals were investigated. In previous studies, i.e. [6], it was shown that average power of tracheal sound and flow has a linear relationship. Thus, at each flow rate, minimizing the mean square error was used to fit a line to the calculated average power by the two mentioned equations:

$$AvgPwr_i = a_P \times F_i + b_P, \quad (4)$$
$$AvgLog_i = a_L \times F_i + b_L, \quad (5)$$

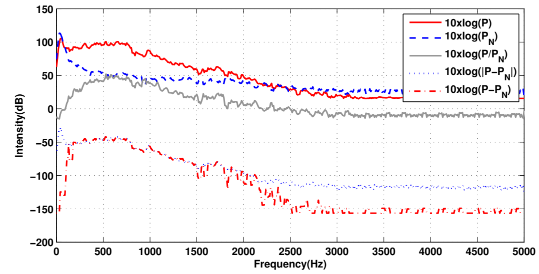
where $AvgPwr_i$ and $AvgLog_i$ represent the average–power and the average–log values of tracheal sound in each segment (i), respectively, and F_i is the amount of airflow in the same segment. For each subject, the coefficients a_P, b_P, a_L and b_L were calculated at different flow rates, and averaged among different subjects for each respiratory phase.

The average values of each coefficient $(a_P, b_P, a_L$ and $b_L)$ at different flow rates were approximated by a line. The normalized mean square error $(Nmse)$ between the average values of each coefficient and their estimated values are calculated as:

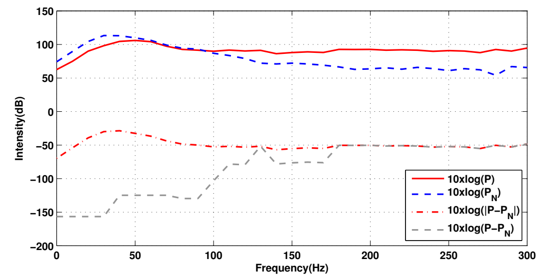
$$Nmse = \frac{E\left[a - \hat{a}\right]^2}{E\left[a\right]^2}. \quad (6)$$

III. Results and Discussion

The PSD of the recorded tracheal sound signal $(10\log(P))$, of a normal adult female subject (section II-C), along with the PSD values of its corresponding noise signals $(10\log(P_N))$ are shown in Fig. 2–a. The third curve shows the results of noise removal by subtracting the power values in dB $(10\log(P/P_N)$, Eq.(2)). The forth curve shows the results after subtracting the power values in Watts and transforming to the logarithmic scale, but without considering the mathematical issues described in section II-A.3 $(10\log(|P - P_N|))$. Comparing these results with those of the previous studies on the frequency components of tracheal sounds, i.e. [7], it can be concluded that the curve representing $10\log(|P - P_N|)$ follows the results of the previous studies, while the $10\log(P/P_N)$ curve is far from the results reported previously. The last curve in Fig. 2–a, $10\log(P - P_N)$, presents the issue addressed in section II-A.3. For deriving the results shown in this curve, the negative values of $P - P_N$ are replaced by a small number of $\epsilon = 2.22\times10^{-16}$. This guarantees the output of the logarithm



(a)



(b)

Fig. 2. PSD of recorded tracheal sound, noise and the tracheal sound after noise reduction a) in the overall frequency range and b) zoom in of the results in the low frequency components.

function will be a real value. Comparison of the results displayed for $\log(|P - P_N|)$ and $\log(P - P_N)$ reveals that they are the same for medium frequency components, where the power values of the recorded sounds are greater than those of the noise. But, at low frequency components, where noise dominates tracheal sound, the results of $\log(|P - P_N|)$ and $\log(P - P_N)$ are different. Since the low frequency components of tracheal sounds are of great importance, the zoomed in of the results at low frequency components are shown in Fig. 2–b. The results of $\log(P - P_N)$ show that we can not extract tracheal sound components below 100 Hz.

In the next experiment, average–power $(AvgPwr)$ and average–log $(AvgLog)$ values of tracheal sound are used to find the relationship between tracheal sound amplitude and airflow at different flow rates. Inspiration and expiration phases are analyzed separately and the average and standard deviation values of the regression coefficients $(a_P, b_P, a_L$ and $b_L)$ are calculated among different subjects (Fig. 3). In order to have a better representation, only half of the standard deviation values are plotted. The results shown in Fig. 3 indicate that the differences between the regression coefficients are more evident at low flow rate and during expiration phase. On the other hand, the regression coefficients results from $AvgPwr$ values seem to fit better to a line, which is the underlying assumption in linear modeling of flow–sound relationship.

A crucial challenge in flow estimation methods is to find a feature of tracheal sound for which its variation at different flow rates is predictable and easy to estimate [8]. If the regression coefficients at different flow rates can be approximated with a line, calibration part of the flow estimation methods will be simplified to a large extent. In this case, the method can be calibrated at one flow rate and adapt itself to the variations of the flow automatically.

To investigate the effect of the two average power calculation methods on this issue, a line is fitted to the average values of each regression coefficient at different flow rates and the error between the linear approximations and the real values is calculated. Fig. 4 presents the normalized mean square error values $(Nmse)$ between the average values of different regression coefficients and their linear approximations. Since the variations of $Nmse$ among different coefficients are large, they are described in logarithmic scale. From these results, it is evident that for all regression coefficients, those resulted from $AvgPwr$ values fit better to a line at different flow rates. These results arise from the underlying acoustical relationship between air flow and the generated sounds power values [3] which will be manipulated when using $AvgLog$ values.

IV. Conclusion

In this paper, the mathematical issues regarding the spectral analysis of sound signals were discussed and the nonlinear effects of logarithm function were examined. Also, the acoustical interpretations corresponding to different mathematical implementations were discussed. As a practical example, respiratory sound analysis was considered and the effects of these arithmetic misconceptions on different applications such as noise removal, average power calculation and flow–sound relationship were explained in detail. The results indicate that some of the most common errors occur in noise reduction. These errors are due to substituting noise reduction by sound detection (SNR) and/or representing the noise power in the very low frequency components instead of the signal power. Also, it was shown that if average of the power components is calculated in the wrong scale, the results do not represent the underlying acoustic characteristics of the sound, and it affects the flow-sound relationships at different flow rates.

References

[1] L. Kinsler, A. Frey, A. Coppens, and J. Sanders, *Fundamentals of acoustics*. USA, New York: John Wiley and Sons Inc., 2005.

[2] M. Golabbakhsh, Z. Moussavi, and M. Aboofazeli, "Respiratory flow estimation from tracheal sound by adaptive filters," in *IEEE-EMBS*, Shanghai, 2005, pp. 4216–4219.

[3] H. Pasterkamp, S. Kraman, and G. Wodicka, "Respiratory sounds, advances beyond the stethoscope," *Am J Respir Crit Care Med*, vol. 156, pp. 974–987, 1997.

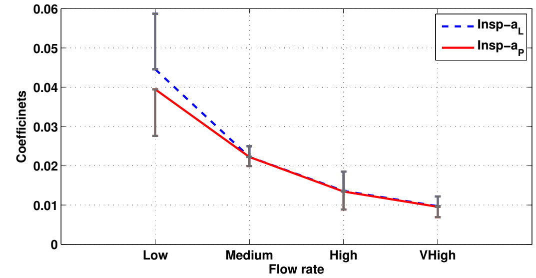
[4] M. Mussel, Y. Nakazano, Y. Miyamoto, S. Okabe, and T. Takishima, "Distinguishing normal and abnormal tracheal breathing sounds by principal component analysis," *Japanese journal of physiology*, vol. 40, No. 5, p. 713721, 1990.

[5] M. Mahagnah and N. Gavriely, "Repeatability of measurements of normal lung sounds," *Am J Respir. Crit. Care Med.*, vol. 149, pp. 477–481, 1994.

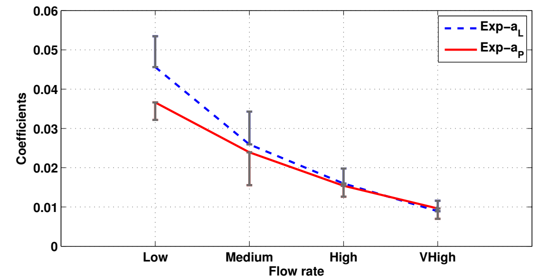
[6] V. Harper, H. Pasterkamp, H. Kiyokawa, and G. Wodicka, "Modeling and measurement of flow effects on tracheal sounds," *IEEE Trans. Biomed. Eng.*, vol. 50, no.1, pp. 1–10, 2003.

[7] I. Sanchez, A. Avital, I. Wong, A. Tal, and H. Pasterkamp, "Acoustic vs. spirometric assessment of bronchial responsiveness to methacholine in children," *Pediatr. Pulmonol.*, vol. 15, No. 1, pp. 28–35, 1993.

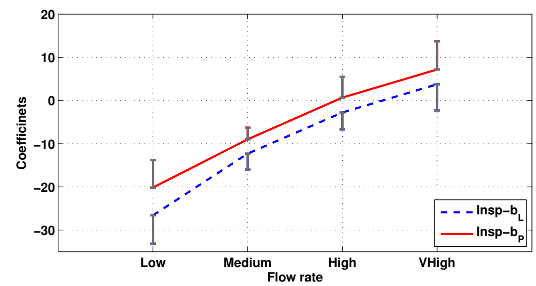
[8] A. Yadollahi and Z. Moussavi, "A robust method for estimating respiratory flow using tracheal sounds entropy," *IEEE Transactions on Biomedical Engineering*, vol. 53, No. 4, pp. 662–668, 2006.



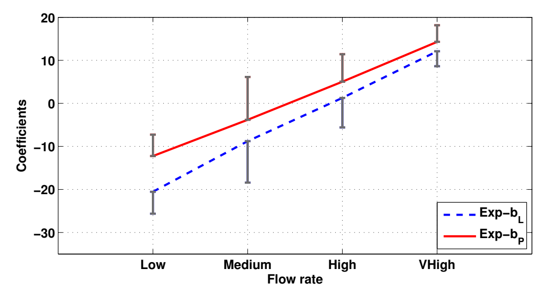
(a)



(b)



(c)



(d)

Fig. 3. Regression coefficients between air flow and $AvgPwr$ and $AvgLog$ values of tracheal sound, a_P and a_L coefficients during a) inspiration, b) expiration, b_P and b_L coefficients during c) inspiration and d) expiration.

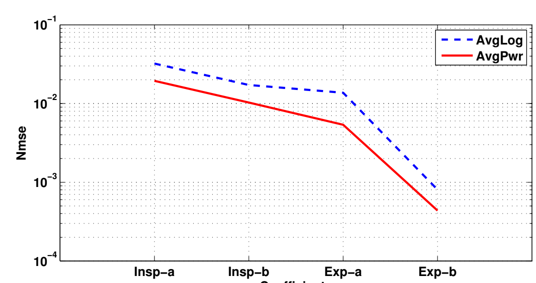


Fig. 4. Normalized mean square error values between different regression coefficient average values and their linear approximations.