Sensors and Materials, Vol. 34, No. 7 (2022) 2723–2734
MYU Tokyo

S & M 3001

Separation of Aortic and Pulmonary Components from Second Heart Sounds without an Assumption of Statistical Independence

Shun Muramatsu,* Seiichi Takamatsu, and Toshihiro Itoh

Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan

(Received November 22, 2021; accepted February 15, 2022)

Keywords: heart sound, A2–P2 splitting interval, independent component analysis, nonlinear transient chirp signal model, demixing vector

A novel algorithm to separate aortic (A2) and pulmonary (P2) components from the second heart sound (S2) without assuming that A2 and P2 are statistically independent, and with optimizing demixing vectors using root-mean-square error (RMSE) between outputs and signal models as cost function is successfully demonstrated. Conventional methods to estimate the A2–P2 splitting interval (SI) based on the separation of A2 and P2 using independent component analysis (ICA) are subject to distortions due to the fact that A2 and P2 are not strictly statistically independent. Therefore, we propose an algorithm to separate A2 and P2 without assuming their independence. In the proposed algorithm, a nonlinear transient chirp signal model is introduced as the proper models of A2 and P2, and the separated sound is optimized to be closest to the A2/P2-like model. To evaluate the proposed algorithm, SI estimation was performed for S2 simulated with 60 common SI patterns. The results show that the proposed algorithm can estimate SI stably regardless of the independence of A2 and P2, and can estimate SI with 95% limits of agreement of −0.305 ± 2.15 ms, which is about 69% smaller as the error range than ICA.

1. Introduction

Cardiovascular disease is the most common cause of death worldwide, accounting for about 32% of all deaths in 2019.[1] Among these diseases, pulmonary arterial hypertension is a frequent and serious complication of several cardiovascular and respiratory diseases, which is difficult to assess noninvasively.[2] Therefore, there is a growing need for a non-invasive method to estimate pulmonary arterial pressure (PAP) accurately and safely. It is known that pulmonary arterial hypertension increases the time interval between the onset of the aortic (A2) and pulmonary (P2) components of the second heart sound (S2) and the dominant frequency of P2.[3] Therefore, the A2–P2 splitting interval (SI) is considered to be useful for the non-invasive estimation of PAP.

Some methods to estimate SI in the time-frequency domain have been proposed.[4–6] Xu *et al.* masked the Wigner–Vile distribution on the time-frequency plane of S2, estimated the

*Corresponding author: e-mail: smuramatsu@g.ecc.u-tokyo.ac.jp
https://doi.org/10.18494/SAM3738

ISSN 0914-4935
https://myukk.org/

instantaneous frequency and amplitude changes of A2 and P2, and reconstructed them from S2 to obtain the waveform of P2. However, this method does not give optimal results when the instantaneous frequencies of A2 and P2 overlap significantly in the time-frequency plane, because S2 is regarded as a monocomponent.[7] Therefore, Vivek *et al*. proposed a method to automatically extract the A2 and P2 components from S2 based on the assumption that A2 and P2 are statistically independent in time.[7] This assumption of mutual independence of A2 and P2 enables the separation of A2 and P2 from S2 using the blind source separation (BSS) technique. There are several attempts to separate A2 and P2 from S2 by applying BSS techniques, especially independent component analysis (ICA).[7–10] However, there is no guarantee that ICA will work properly, because A2 and P2 have overlapping timing, and they are not considered to be strictly statistically independent.

The purpose of this study is to propose a new algorithm to extract A2 and P2 from S2 without an assumption that A2 and P2 are statistically independent. The proposed algorithm separates S2 into the most A2/P2-like sounds, optimizing it to be close to the A2 and P2 models.

2. Proposed Algorithm

2.1 Overview of proposed algorithm

Figure 1 shows an overview of our proposed algorithm. The proposed method outputs the A2 and P2 by optimizing the correct demixing vector and multiplying it by the observed S2. This approach is similar to that of the demixing matrix in ICA. ICA uses the kurtosis indicating a degree of independence as a cost function for optimizing the demixing matrix.[11] In contrast, the proposed method introduces signal models for A2 and P2, and uses the root-mean-square error (RMSE) between the models and the output as a cost function.

There are several attempts to model A2 and P2.[12–14] In this study, we adopt the nonlinear transient chirp signal model.[12] According to this model, S2 is expressed as Eqs. (1)–(3).

$$\boldsymbol{S2}(t) = \boldsymbol{A}_{\mathrm{A2}}(t)\sin(\boldsymbol{\varphi}_{\mathrm{A2}}(t)) + \boldsymbol{A}_{\mathrm{P2}}(t-t_0)\sin(\boldsymbol{\varphi}_{\mathrm{P2}}(t)) + \boldsymbol{n}(t) \tag{1}$$

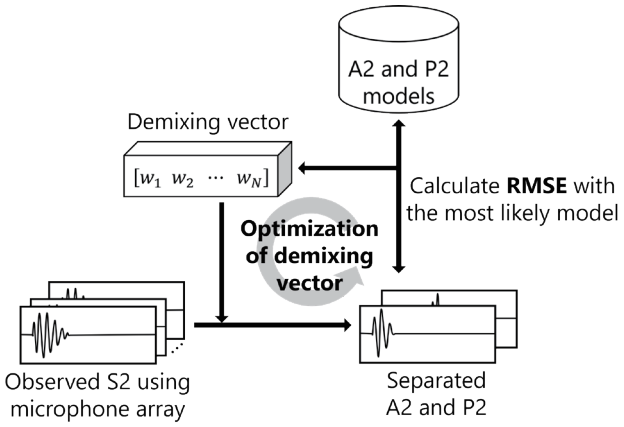

Fig. 1. Overview of proposed algorithm.

$$\boldsymbol{A}(t) = a_0\left(1 - e^{-\frac{t}{8}}\right)e^{-\frac{t}{16}}\sin\frac{\pi t}{60} \tag{2}$$

$$\boldsymbol{\varphi}(t) = 2\pi\left(f_1 t + 2f_2 t^{\frac{1}{2}}\right) \tag{3}$$

The first term of $\boldsymbol{S2}(t)$ represents the A2 component and the second term represents the P2 component. $\boldsymbol{A}(t)$ and $\boldsymbol{\varphi}(t)$ are the time variations of the amplitude and frequency of A2 and P2 corresponding to the subscripts. t_0 is the time of occurrence of A2 and P2. $\boldsymbol{n}(t)$ is the additive white Gaussian noise. These equations include the two parameters as characteristic frequencies (f_1 and f_2). These express individual differences in the waveform. Since a single valvular sound is a simple sound such as a collision sound associated with the closing of a valve, these two parameters can absorb individual differences.

Figure 2 shows the waveforms of A2 and P2 generated by Eqs. (1)–(3) with the following parameters also shown in Table 1. a_0 is a parameter used to adjust the amplitude, so it is set to 1.0 here. $f_1 = 24, f_2 = 226$ for A2 and $f_1 = 22, f_2 = 178$ for P2 based on previous studies.[12] In practice, these two components are mixed by some mixing matrix and observed as S2 on the chest wall using a stethoscope, for example.

2.2 Signal processing flow

Figure 3 shows the signal process flow of the proposed algorithm. $\boldsymbol{S2}_a$ and $\boldsymbol{S2}_b$ are two observed *S*2s. The A2/P2 model list includes 12000 A2/P2 models generated using Eqs. (1)–(3).

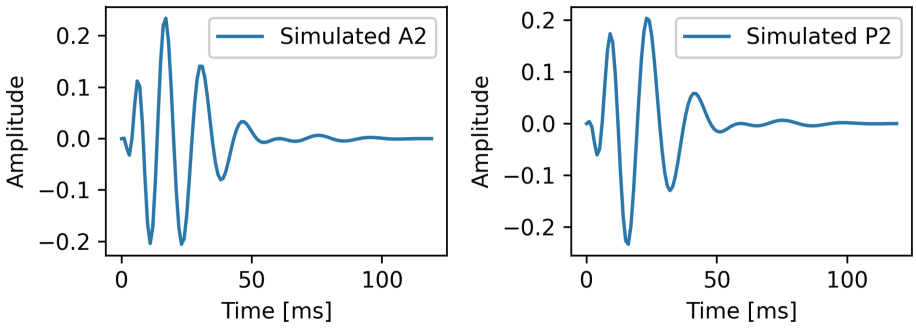

Fig. 2. (Color online) Simulated waveforms of A2 and P2.

Table 1
Parameters of nonlinear transient chirp signal model for A2, P2, and A2/P2 model list.

	t_0 (ms)	a_0	f_1 (Hz)	f_2 (Hz)
A2	0	1	24	226
P2	0–59	1	22	178
A2/P2 model list	0–59	1	20	100–300

These are fixed f_1 to 20 Hz, swept f_2 from 100 Hz to 300 Hz, and swept t_0 from 0 to 59 ms as shown in Table 1. Originally, f_1 is also variable, but since the same waveform can be expressed by only changing f_2 as when f_1 is changed, f_1 is fixed this time. First, the demixing vector $\boldsymbol{w}$ is optimized so that the RMSE between $\boldsymbol{w}[\boldsymbol{S2}_\mathrm{a}\ \boldsymbol{S2}_\mathrm{b}]^\mathrm{T}$ and the model is minimized. This is done for all 12000 models.

Figure 4 shows an example of the RMSE contour calculated for all models. Here, a low RMSE indicates that the observed *S*2 is likely to contain that model. This is because $\boldsymbol{w}[\boldsymbol{S2}_\mathrm{a}\ \boldsymbol{S2}_\mathrm{b}]^\mathrm{T}$ with the correct demixing vector represents one true source signal. Therefore, we select the two models with the lowest RMSE as the most likely A2 and P2 models, as shown by the arrows in Fig. 4. This example is for *S*2, a mixture of A2 with a t_0 of 0 ms and P2 with a t_0 of 20 ms, and the correct t_0 is selected. There may be some errors for some *S*2; these errors are due to the fact that f_1 is fixed, which is not a problem for the subsequent separation process in the proposed algorithm.

Finally, $\boldsymbol{w}_\mathrm{A2}$ and $\boldsymbol{w}_\mathrm{P2}$ are optimized again to minimize the RMSE between the separated sound and the selected model, and output $\boldsymbol{y}_\mathrm{A2}$ and $\boldsymbol{y}_\mathrm{P2}$ corresponding to the estimated A2 and P2.

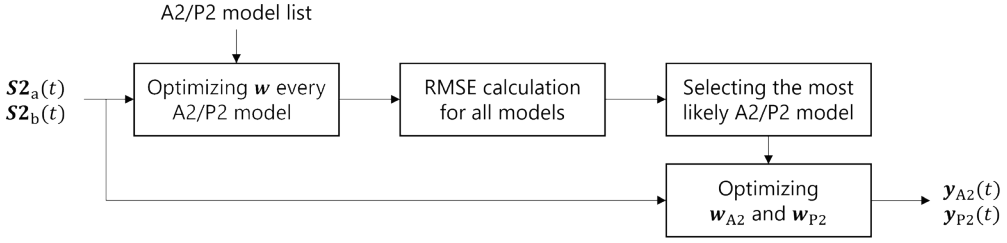

Fig. 3. Signal process flow of proposed algorithm.

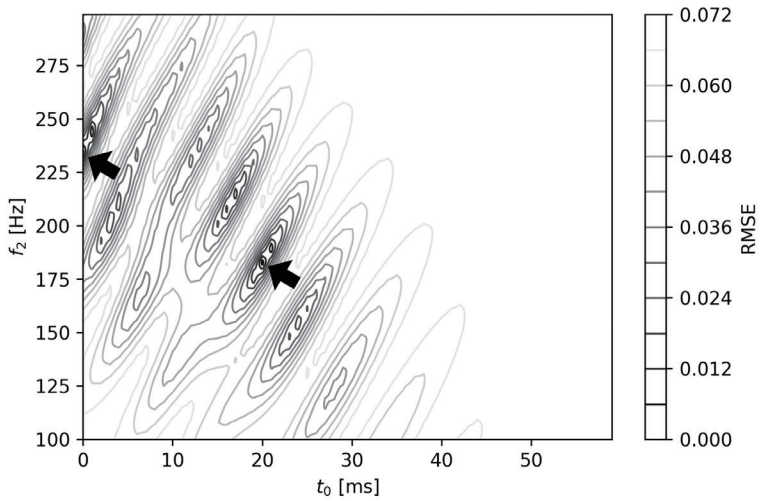

Fig. 4. Example of RMSE contour for every A2/P2 model. The two arrows point to the most A2/P2-like models.

3. Evaluation Method

3.1 Separation of A2 and P2 from S2

The evaluation experiment is conducted with simulated S2 sounds. Figure 5 shows an example of two S2s simulated under the mixing conditions shown in Fig. 6, where the aortic and pulmonary valves are 50 mm apart and the distance between the two stethoscopes on the chest wall is also 50 mm. This spatial condition is not strictly necessary; we just needed to determine this condition to generate the mixing matrix for simulating S2. It is only important that the two stethoscopes are not in the positions where the observed sounds are the same; in other words, the mixing matrix is invertible. This mixing is instantaneous, which does not take into account the reverberation and time delay in the actual propagation path from the heart valve to the chest wall. It has been reported that the source separation algorithm for instantaneous mixing can be extended to the time-frequency domain to deal with reverberation and delay.[15] For simplicity, we discuss the proposed algorithm under the condition of instantaneous mixing.

Sixty pairs of S2s with different SIs were simulated. The sampling frequency is 1000 Hz. As an example, the waveforms of the simulated S2 with SIs of 0, 10, 20, and 30 ms are shown in Fig. 7. When the SI is less than about 20 ms, it is difficult to estimate the SI from the waveform

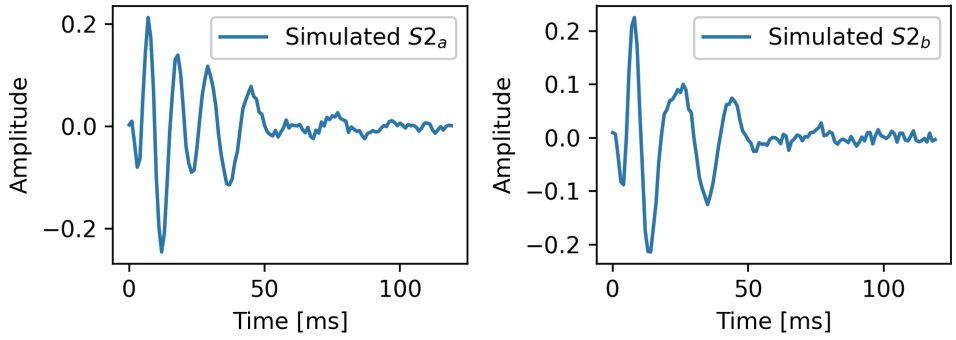

Fig. 5. (Color online) Two examples of simulated waveform of S2.

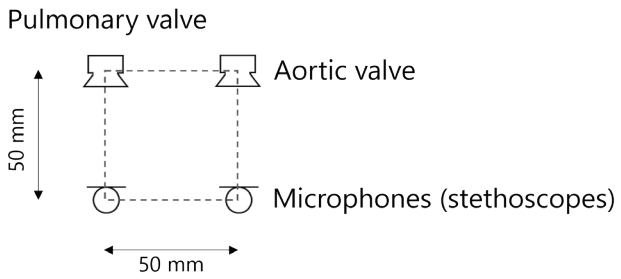

Fig. 6. Spatial conditions during A2 and P2 mixing simulations.

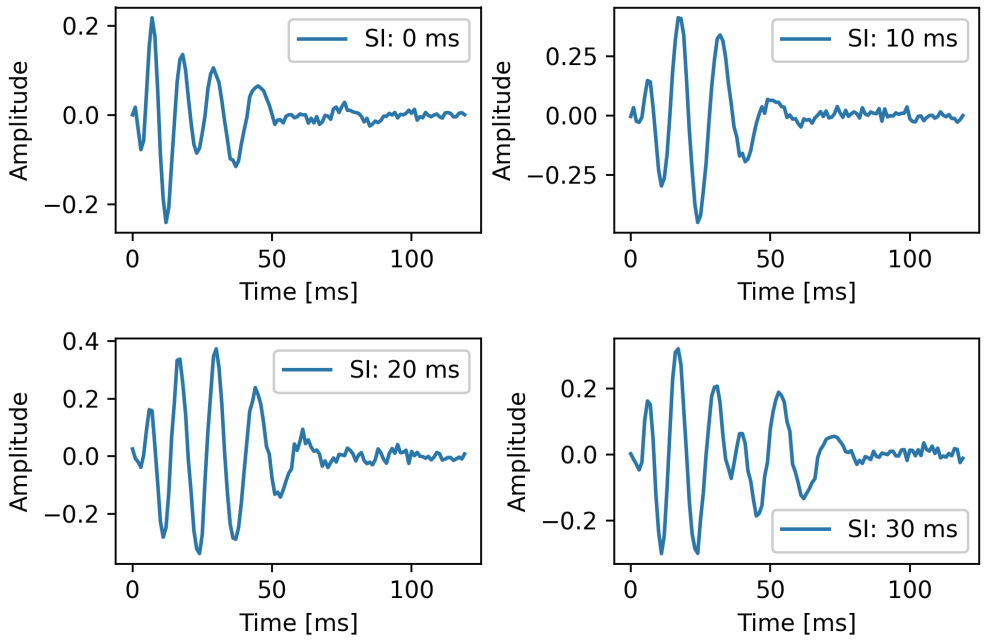

Fig. 7. (Color online) Examples of S2 simulated with four different SIs.

or even to detect the split between A2 and P2. The separation of the proposed algorithm is evaluated by separating 60 pairs of S2 into A2/P2 using both ICA and the proposed algorithm and calculating the RMSE between the separated waveform and the true waveform.

3.2 Estimation of SI between A2 and P2

After separating A2 and P2 from these simulated S2, the SI is estimated. The time-centroid-based method[7] is then used to estimate the SI. The time centroid c_x of N samples of a signal $\boldsymbol{x}(t)$ is defined as

$$c_x = \sum_{i=1}^{N} iT \,\mathrm{abs}(\boldsymbol{x}(iT)) \Big/ \sum_{i=1}^{N} \mathrm{abs}(\boldsymbol{x}(iT)) \,. \tag{4}$$

T is the sampling time interval. In other words, when c_{A2} and c_{P2} are the time centroids of the separated A2 and P2, respectively, the SI, t_s, is calculated as $t_s = c_{P2} - c_{A2}$. The accuracy of the SI estimation was evaluated by calculating the SI from the separated signals using ICA and the proposed algorithm and comparing it with the true value. Note that the FastICA algorithm in the python library scikit-learn is used for ICA, and the Broyden–Fletcher–Goldfarb–Shanno algorithm was used for optimization in the proposed algorithm.

4. Results and Discussion

4.1 Separation of A2 and P2

First, conventional ICA attempted to separate the simulated S2 into A2 and P2. To evaluate the performance, RMSE between the separated A2/P2 and the true A2/P2 is calculated; the lower the RMSE, the better the separation performance. Thirty trials were conducted for all S2s, and 1800 data were obtained. Figure 8 shows the average value of RMSE for each SI. When the SI is 27 ms or higher, RMSE is stable at about 0.02, which indicates successful separation. On the other hand, when the SI is less than 26 ms, RMSE is not stable and A2/P2 is not separated successfully with most of the SIs. This is due to the distortion caused by trying to force two signals that are not independent from each other to become independent. As an example, Fig. 9 shows separated sounds that failed to separate when the SI was 0 ms. ICA does not inherently maintain the scale of the waveform. Therefore, we used the back projection[16] to match the scale of the observed sound on the chosen microphone. A2 has a high distortion, and P2 has a reduced amplitude. These waveforms have some distortions caused by the non-independence of A2 and P2.

The same process was applied to the proposed algorithm. As an example, Fig. 10 shows the separated sounds when SI is 0 ms. Compared with the ICA separation shown in Fig. 9, the separated A2 and P2 almost have no distortion. The box plots of the RMSE obtained by the two methods are shown in Fig. 11. In this figure, the results when the SI is 26 ms or less are plotted,

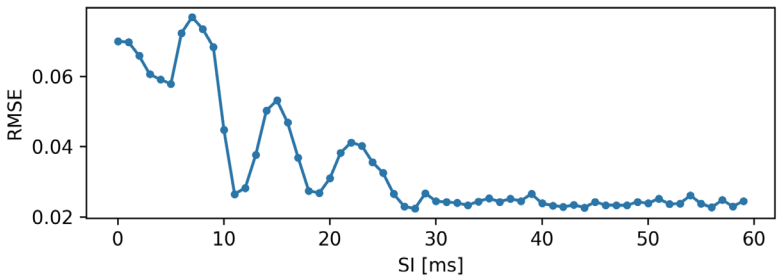

Fig. 8. (Color online) RMSE for A2/P2 separation using ICA against every SI.

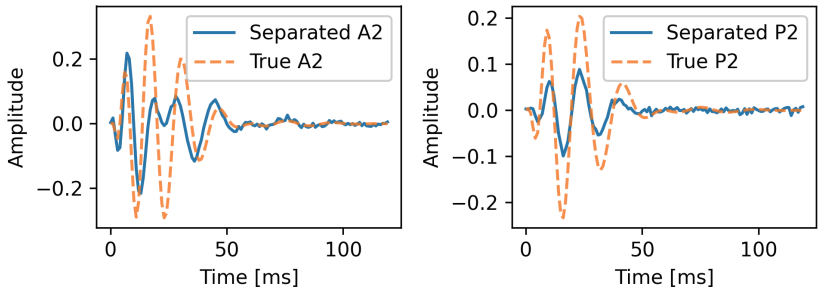

Fig. 9. (Color online) Waveforms of separated A2 and P2 using ICA.

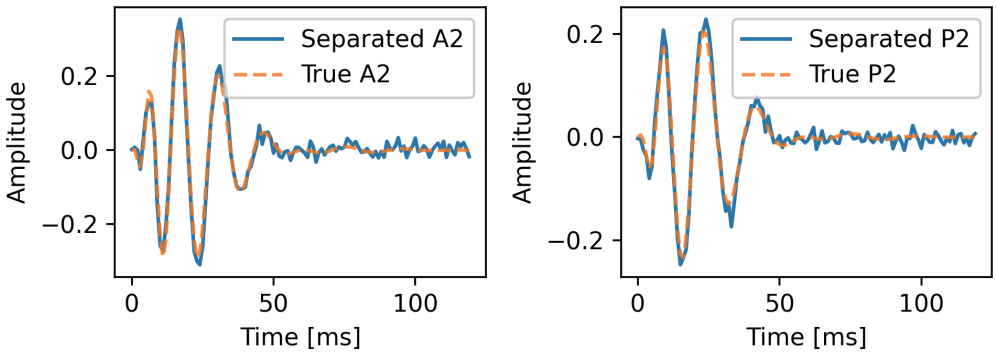

Fig. 10. (Color online) Waveforms of separated A2 and P2 using the proposed algorithm.

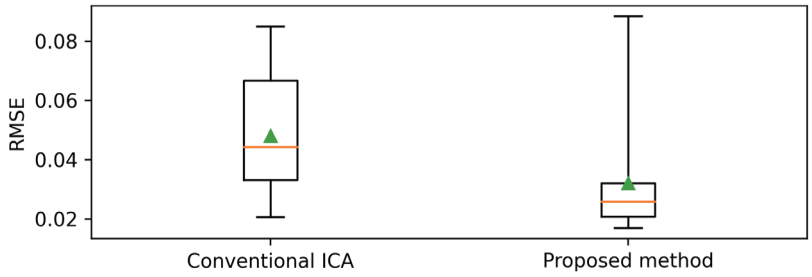

Fig. 11. (Color online) Box plots of RMSE as separation performance of ICA and the proposed algorithm.

because both methods obtained good RMSE when the SI is 27 ms or more. The mean value of the proposed algorithm is 0.032, which is 0.017 smaller than the ICA value of 0.049, indicating high accuracy. The interquartile range of the proposed algorithm is 0.011, which is 0.024 smaller than that of ICA (0.035), indicating stable results with small variation.

4.2 Estimation of SI between A2 and P2

The SI was estimated from the A2 and P2 waveforms separated in the previous section using the time-centroid method. We discuss only the results when the SI is less than 20 ms, in which it is difficult to detect the A2–P2 splitting visually. Figure 12(a) shows the relationship between the true SI and the SI estimated by ICA. The slope of the regression line denoted by the dashed line is 0.659, and the correlation coefficient is 0.844. The closer these values are to 1, the more correctly the SI is estimated. Figure 12(a) shows that the SI is not estimated well, especially in the region where the SI is close to 0 ms. On the other hand, Fig. 12(b) shows the result of the proposed algorithm. The slope of the regression line and the correlation coefficient are 0.959 and 0.982, respectively, indicating that the proposed algorithm can estimate the SI with higher accuracy than ICA.

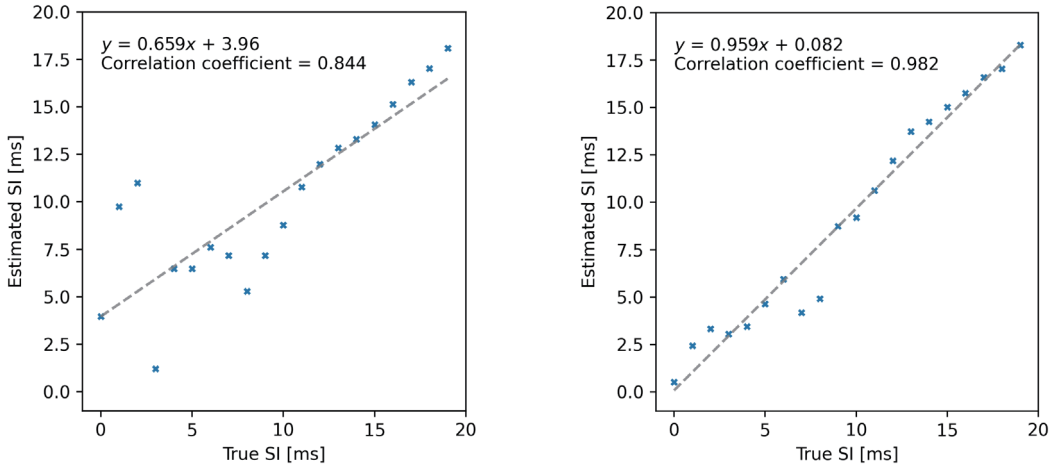

Fig. 12. (Color online) Comparison of estimated SI with true SI using (a) ICA and (b) the proposed algorithm. The dashed line is the linear regression line.

Figure 13 shows the Bland–Altman analysis[17] of these results, which plots the difference between the estimated SI and true SI values against the mean of the estimated and true SI values. Figure 13(a) shows the result of ICA. In particular, the estimation accuracy is very low and unstable when the mean of the estimated and true SI is between 0 ms and about 8 ms. The 95% limit of agreement (LOA), denoted by the dashed line, is −0.649 ± 6.85 ms. On the other hand, Fig. 13(b) shows the result of the proposed algorithm: the estimation accuracy is stable regardless of SI, and LOA is −0.305 ± 2.15 ms, which is about 9.40 ms smaller than that of ICA. These indicate that the proposed algorithm can estimate SI with less error than ICA.

Although there is no general required accuracy about the SI estimation yet, it can be considered as follows. The only method that can directly measure PAP is pulmonary artery catheterization, and it is invasive. The only practical method of non-invasive but restrictive estimation of PAP is Doppler echocardiography, and its standard error of the estimate (SEE) against the catheterization is 8 mmHg.[18] On the other hand, our proposed method by SI estimation can calculate PAP (as mean of systole and diastole) using Eqs. (5) and (6).[2,19]

$$P_{\mathrm{PAP}} = -13.7 + 19.2\sqrt{I_{\mathrm{NSI}}} \tag{5}$$

$$I_{\mathrm{NSI}} = \frac{I_{\mathrm{SI}} R_{\mathrm{HR}}}{600} \tag{6}$$

Here, P_{PAP} is PAP (mmHg), I_{NSI} is normalized SI (ms), I_{SI} is SI (ms), and R_{HR} is heart rate (bpm). From these equations, 8 mmHg as SEE can be converted to LOA. The converted LOA is *mean* ±3.48 ms (*mean* is not important here) assuming a heart rate of 60. The LOA of our proposed method is ± 2.15 ms, which is lower than that of Doppler echocardiography (±3.48 ms).

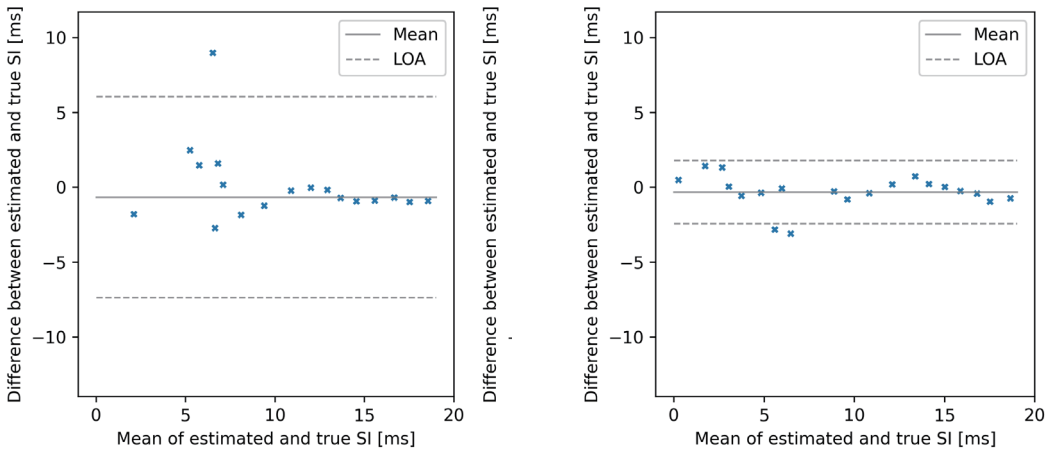

Fig. 13. (Color online) Bland–Altman analysis of the estimated SI using (a) ICA and (b) the proposed algorithm. The solid line indicates the mean of the difference between the estimated and true SIs. The dashed lines indicate LOA.

It indicates that the proposed method is practical. However, it is important to note that these results are based on simulations, and it is necessary to continue experiments on actual heart sounds and improvements on the algorithm.

4.3 Limitations

Figure 11 shows the separation performance of the proposed algorithm in terms of RMSE. The overall RMSE is low, indicating that the proposed algorithm has stable and high separation performance compared with ICA. On the other hand, the RMSE of some samples exceeded 0.08. The reason for this is that the proposed algorithm mistakenly selected a model that has a similar waveform, but the generation timing is off by one wavelength. However, Fig. 13(b) shows that the SI estimation result of the proposed algorithm is very stable. The reason for this is that even if a wrong waveform is mistakenly selected, the proposed algorithm does not distort the A2 and P2 waveforms significantly. For that reason, the time-centroid method is performed well with some wrong selection. As a result, the proposed algorithm can estimate the SI stably with LOA of -0.305 ± 2.15 ms, which is 9.40 ms smaller than the ICA.

The proposed algorithm needs to select the two models that are most likely to be A2 and P2, as shown by two arrows in Fig. 4. This time, this selection process worked well. However, if, for example, A2 and P2 have the same parameters (f_1 and f_2) and the SI is 0 ms, i.e., A2 and P2 are not split, the minima to be selected may overlap. Then the overlapping minima will be selected as the A2 model, and then the model with the second lowest RMSE will be selected as the P2 model, even though it is not the correct model. In other words, the model selection process of the proposed algorithm does not work in this situation. Therefore, we are trying to solve this problem by determining that the SI is close to 0 ms by considering the respiratory variability of the SI.

5. Conclusions

A novel algorithm to separate A2 and P2 from S2 with optimizing demixing vectors using RMSE between outputs and A2/P2 signal models as cost function is successfully demonstrated. This proposed algorithm does not require an assumption that A2 and P2 are statistically independent. In this algorithm, the nonlinear transient chirp signal model is introduced as a model of A2 and P2. First, this algorithm seeks two models that are most close to multiplication of the observed sound and the optimized demixing vector. Then, by optimizing the separated sound to be close to the sought models, the proposed algorithm can separate S2 into the most A2/P2-like sounds. To evaluate the proposed algorithm, 120 types of S2 with different SIs are simulated, and the separation of A2/P2 and the estimation of SI are performed using the proposed algorithm. As a result, in the waveform separation, the mean value of RMSE, which indicates the magnitude of waveform distortion, was reduced by about 35% and the interquartile range by about 69% using the proposed algorithm compared with the conventional ICA algorithm. In the estimation of SI, the correlation coefficient between the SI estimated using the proposed algorithm and the true SI is 0.982. Furthermore, the proposed algorithm can be used to estimate the SI with an error of -0.305 ± 2.15 ms as LOA. This error range is about 69% smaller than that obtained by the conventional method (-0.649 ± 6.85 ms). These mean that the proposed algorithm can be used to estimate SI stably regardless of the independence of A2 and P2. By estimating PAP from SI calculated by this method, we can establish an accurate and safe non-invasive evaluation method for pulmonary arterial hypertension.

References

1 World Health Organization: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed November 2021).
2 J. Xu, L. G. Durand, and P. Pibarot: Heart **88** (2002) 76. https://doi.org/10.1136/heart.88.1.76
3 B. Popov, G. Sierra, L. G. Durand, J. Xu, P. Pibarot, R. Agarwal, and V. Lanzo: Proc. 26th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (2004) 921. https://doi.org/10.1109/IEMBS.2004.1403310
4 J. Xu, L. G. Durand, and P. Pibarot: IEEE Trans. Biomed. Eng. **47** (2000) 1328. https://doi.org/10.1109/10.871405
5 H. Luo, P. Westphal, M. Shahmohammadi, L. I. B. Heckman, M. Kuiper, R. N. Cornelussen, T. Delhaas, and F. W. Prinzen: Physiol. Rep. **9** (2021) e14687. https://doi.org/10.14814/phy2.14687
6 A. Djebbari and F. Bereksi-Reguig: Biomed. Eng. Online **12** (2013) 1475. https://doi.org/10.1186/1475-925X-12-37
7 V. Nigam and R. Priemer: Physiol. Meas. **27** (2006) 553. https://doi.org/10.1088/0967-3334/27/7/001
8 P. Comon: Signal Process. **36** (1994) 287. https://doi.org/10.1016/0165-1684(94)90029-9
9 L. Chen, S. F. Wu, Y. Xu, W. D. Lyman, and G. Kapur: J. Theor. Comput. Acoust. **26** (2018) 1750035. https://doi.org/10.1142/S2591728517500359
10 C. Xiefeng, D. Shicheng, L. Yunyi, S. Xu, and L. Jia: Proc. 2018 Chinese Control and Decision Conf. (2018) 5573. https://doi.org/10.1109/CCDC.2018.8408103
11 A. Hyvarinen and E. Oja: Neural Networks **13** (2000) 411. https://doi.org/10.1016/S0893-6080(00)00026-5
12 J. Xu, L. G. Durand, and P. Pibarot: IEEE Trans. Biomed. Eng. **48** (2001) 277. https://doi.org/10.1109/10.914790
13 A. Baykal, Y. Ziya, and H. Koymen: IEEE Trans. Biomed. Eng. **42** (1995) 358. https://doi.org/10.1109/10.376129
14 X. Zhang, L. G. Durand, L. Senhadji, H. C. Lee, and J. L. Coatrieux: IEEE Trans. Biomed. Eng. **45** (1998) 962. https://doi.org/10.1109/10.704865
15 H. Sawada, N. Ono, H. Kameoka, D. Kitamura, and H. Saruwatari: APSIPA Trans. Signal Inf. Process. **8** (2019) e12. https://doi.org/10.1017/ATSIP.2019.5
16 D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari: IEEE Trans. Audio Speech Lang. Process. **24** (2016) 1626. https://doi.org/10.1109/TASLP.2016.2577880

17 J. M. Bland and D. Altman: The Lancet **327** (1986) 307. https://doi.org/10.1016/S0140-6736(86)90837-8
18 P. G. Yock and R. L. Popp: Circulation **70** (1984) 657. https://doi.org/10.1161/01.CIR.70.4.657
19 T. S. Leng, P. R. White, J. Cook, W. B. Collis, E. Brown, and A. P. Salmon: IEE Proc.: Sci. Meas. Technol. **145** (1998) 285. https://doi.org/10.1049/ip-smt:19982326