2022 International Workshop on Acoustic Signal Enhancement (IWAENC) | 978-1-6654-6867-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/IWAENC53105.2022.9914748

PHASE ERROR ANALYSIS FOR FIRST-ORDER LINEAR DIFFERENTIAL MICROPHONE ARRAYS

Longfei Yan[1,2*], *Weilong Huang*[3], *W. Bastiaan Kleijn*[1], *Thushara D. Abhayapala*[2]

[1]School of Engineering and Computer Science, Victoria University of Wellington,
[2]College of Engineering and Computer Science, Australian National University,
[3]DingTalk Hummingbird Audio Lab, Alibaba Group

ABSTRACT

First-order Linear Differential Microphone Arrays (LDMAs) are sensitive to sensor imperfections such as phase errors. This paper analyses the impacts of both bounded and unbounded phase errors on first-order LDMAs. We propose a tolerance threshold that allows bounded phase errors to take various values. Moreover, White Noise Gain (WNG) thresholds for preventing mainlobe misorientation are obtained. Our work provides guidance for the design of robust first-order LDMAs.

Index Terms— differential microphone array, phase error, white noise gain, mainlobe misorientation

1. INTRODUCTION

Microphone arrays are indispensable components of many hands-free communication systems and speech recognition systems in adverse environments [1–5]. Recently, LDMAs have attracted a significant amount of interest, since they possess a few advantages over traditional methods. Firstly, they can construct relatively frequency-invariant beampatterns, and hence are appropriate for speech signal processing. Secondly, they facilitate large Directivity Factors (DFs) with small and compact apertures [6, 7].

It is well known that LDMAs, especially those of high order, suffer from white noise amplification and struggle to have high WNG values [8, 9]. As a result, LDMAs are sensitive to sensor imperfections that worsen the robustness and reduce WNG values [10, 11]. First-order LDMAs are commonly used to mitigate the influence of sensor imperfections.

Many works have investigated how sensor imperfections affect first-order LDMAs [12–14]. In [12], DF and Front-to-Back Ratio (FBR) lower bounds are optimised with respect to sensor imperfections, but WNG optimisation is lacking. Recent works [13, 14] show that sensor phase errors have a critical influence on the mainlobe orientation of first-order LDMAs pointing towards the endfire direction. However, the tolerance of phase errors claimed in [13, 14] does not reflect practical scenarios by assuming that all microphones have nearly identical values of bounded phase errors, which are close to 0. Moreover, there is necessity to investigate unbounded phase errors that can be arbitrarily large. Although they have been overlooked so far, unbounded phase errors naturally arise in scenarios like wireless acoustic sensor networks where device synchronisation is challenging [15–18].

In this work, we propose a tolerance threshold for bounded phase errors without the assumption that microphones should have similar phase errors. We also derive WNG thresholds for preventing the mainlobe misorientation due to bounded and unbounded phase errors.

*This work was supported by Alibaba Group through Alibaba Research Intern Program.

2. BACKGROUND

2.1. Signal Model

Consider a uniform linear array of M omnidirectional microphones with inter-microphone spacing δ. The steering vector of this array is defined as [19]

$$\mathbf{d}(\omega, \cos\theta) = [1 \; e^{-\jmath\omega\tau_0\cos\theta} \; \cdots \; e^{-\jmath(M-1)\omega\tau_0\cos\theta}]^T, \quad (1)$$

where $\omega = 2\pi f$ is the angular frequency and f is the temporal frequency, θ is the direction of arrival of the source signal from the array axis, $\jmath$ is $\sqrt{-1}$, c is the speed of sound in air and $\tau_0 = \delta/c$.

Beampattern characterizes the input-output behaviour of microphone arrays in beamforming [20]. It is defined as

$$\mathcal{B}_M(\omega, \theta) = [\mathbf{d}(\omega, \cos\theta)]^H \mathbf{h}(\omega) \quad (2)$$

$$= \sum_{m=1}^{M} H_m(\omega) e^{\jmath(m-1)\omega\tau_0\cos\theta} \quad (3)$$

where $\mathbf{h}(\omega) = [H_1(\omega), H_2(\omega), \cdots, H_M(\omega)]^T$ is composed of the filter coefficients. With distortionless constraint, $\mathcal{B}_M[\omega, \theta]$ has the property that

$$\mathcal{B}_M(\omega, \theta) \begin{cases} = 1, & \theta = \theta_d \\ < 1, & \theta \neq \theta_d \end{cases}, \quad (4)$$

where θ_d is the desired look direction of the beamformer.

2.2. Metrics of First-Order LDMA

By approximating the exponential term in (3) with Taylor series, we get $\mathcal{B}_{M,1}(\theta)$, the beampattern of the first-order LDMA with angle θ:

$$\mathcal{B}_{M,1}(\theta) \approx \sum_{m=1}^{M} H_m(\omega) \sum_{n=0}^{1} \frac{1}{n!}[\jmath(m-1)\omega\tau_0 \cos\theta]^n \quad (5)$$

$$= \sum_{m=1}^{M} H_m(\omega) + \jmath\omega\tau_0 \cos\theta \sum_{m=1}^{M}(m-1)H_m(\omega) \quad (6)$$

$$= a_{1,0} + a_{1,1}\cos\theta. \quad (7)$$

where $a_{1,0}$ and $a_{1,1}$ are the real coefficients of $\mathcal{B}_{M,1}(\theta)$.

From (6) and (7), we can obtain the values for H_m's by following the least-norm principle:

$$H_m(\omega) = \frac{6(m-1)a_{1,1}}{\jmath\omega\tau_0 M(M-1)(2M-1)},\ m = 2, 3, \cdots, M, \quad (8)$$

and

$$H_1(\omega) = a_{1,0} - \sum_{m=2}^{M} H_m(\omega) = a_{1,0} - \frac{3a_{1,1}}{\jmath\omega\tau_0(2M-1)}. \quad (9)$$

The mainlobe orientation θ_{main} is defined as

$$|\mathcal{B}_{M,1}(\theta_{\text{main}})|^2 = \max_{\theta} |\mathcal{B}_{M,1}(\theta)|^2. \quad (10)$$

Generally, $\theta_{\text{main}} = \theta_d$ and $|\mathcal{B}_{M,1}(\theta_{\text{main}})|^2 = 1$.

One common metric to measure the robustness of microphone arrays is by using WNG [21]. It is defined as

$$\text{WNG}[\mathbf{h}(\omega)] = \frac{|\mathcal{B}_{M,1}(\theta_d)|^2}{[\mathbf{h}(\omega)]^H \mathbf{h}(\omega)}. \quad (11)$$

Another commonly used metric is DF:

$$\text{DF}[\mathbf{h}(\omega)] = \frac{|\mathcal{B}_{M,1}(\theta_d)|^2}{\frac{1}{2}\int_0^{\pi} |\mathcal{B}_{M,1}(\theta)|^2 \sin\theta d\theta}. \quad (12)$$

We assume that the source signal is from the endfire direction, i.e., $\theta_d = 0°$. This implies that (7) can be rewritten as $a_{1,0} + a_{1,1} = 1$. We also assume $\delta \ll \lambda$, which means the spacing between microphones should be much smaller than the wavelength. We consider the source signal to be far-field. The reverberation effect is not included in this work.

3. ANALYSIS OF PHASE ERRORS

The steering vector with phase errors can be expressed as

$$\mathbf{d}^{(p)}(\omega, \cos\theta) = [e^{-\jmath\psi_1(\omega)}\ e^{-\jmath\omega\tau_0\cos\theta - \jmath\psi_2(\omega)}\ \cdots\ e^{-\jmath(M-1)\omega\tau_0\cos\theta - \jmath\psi_M(\omega)}]^T, \quad (13)$$

where $\psi_m(\omega)$ is the phase error of the mth microphone at the frequency of ω.

Bounded phase errors generally appear in Analogue-to-Digital Converters (ADCs) used for microphone arrays. They are analysed in detail in Section 3.1. Large phase errors that are unbounded do occur in scenarios like wireless acoustic sensor networks. They are also discussed in Section 3.2.

3.1. Analysis of Bounded Phase Errors

We assume $|\psi_m(\omega)| \leq \epsilon_p$, where ϵ_p is the non-negative phase error bound close to 0. From (13), we can utilize Taylor approximation to derive the beampattern of the first-order LDMA with phase errors as

$$\mathcal{B}^{(p)}_{M,1}(\theta) = [\mathbf{d}^{(p)}(\omega, \cos\theta)]^H \mathbf{h}(\omega) \quad (14)$$

$$\approx a_{1,0} + a_{1,1}\cos\theta + \jmath a_{1,0}\psi_1(\omega) - \frac{3a_{1,1}}{\omega\tau_0(2M-1)} \times \left[\psi_1(\omega) - \frac{2\sum_{m=2}^{M}(m-1)\psi_m(\omega)}{M(M-1)}\right]. \quad (15)$$

The beampattern coefficients distorted by phase errors can be observed from (15), which are

$$a^{(p)}_{1,0} = a_{1,0} + \jmath a_{1,0}\psi_1(\omega) - \frac{3a_{1,1}}{\omega\tau_0(2M-1)} \times \left[\psi_1(\omega) - \frac{2\sum_{m=2}^{M}(m-1)\psi_m(\omega)}{M(M-1)}\right], \quad (16)$$

and $a^{(p)}_{1,1} = a_{1,1}$. To investigate whether phase errors will affect θ_{main}, we take the derivative of $|\mathcal{B}^{(p)}_{M,1}(\theta)|^2$ with respect to θ:

$$\frac{d|\mathcal{B}^{(p)}_{M,1}(\theta)|^2}{d\theta} = -2a^{(p)}_{1,1}\sin\theta \mathcal{B}^{(p)}_{M,1}(\theta). \quad (17)$$

This shows that $\frac{d|\mathcal{B}^{(p)}_{M,1}(\theta)|^2}{d\theta}$ has extrema when $\sin\theta = 0$ or $\mathcal{B}^{(p)}_{M,1}(\theta) = 0$. It is obvious that $\mathcal{B}^{(p)}_{M,1}(\theta) = 0$ implies minima. Since θ_{main} is related to the maxima, we only need to consider $\sin\theta = 0$, where $\theta = 0°$ or $180°$. If $\theta_{\text{main}} = 0°$, phase errors have no effect on θ_{main}. Otherwise, mainlobe misorientation happens when $\theta_{\text{main}} \neq \theta_d$.

Mainlobe misorientation can severely degrade the performance of LDMAs [14]. It follows that the condition to avoid mainlobe misorientation caused by phase errors is $|\mathcal{B}^{(p)}_{M,1}(0°)|^2 \geq |\mathcal{B}^{(p)}_{M,1}(180°)|^2$. We first examine

$$|\mathcal{B}^{(p)}_{M,1}(0°)|^2 - |\mathcal{B}^{(p)}_{M,1}(180°)|^2 = 4a_{1,1}\left(a_{1,0} + \frac{-6\Phi_T}{\omega\tau_0 M(M-1)(2M-1)}a_{1,1}\right), \quad (18)$$

where

$$\Phi_T = \sum_{m=2}^{M}(m-1)(\psi_1(\omega) - \psi_m(\omega)). \tag{19}$$

To satisfy $|\mathcal{B}_{M,1}^{(p)}(0^\circ)|^2 - |\mathcal{B}_{M,1}^{(p)}(180^\circ)|^2 \geq 0$, we get $\Phi_T \leq \zeta$, where ζ is the tolerance threshold for phase errors:

$$\zeta = \frac{M(M-1)(2M-1)\omega\tau_0 a_{1,0}}{6(1-a_{1,0})}. \tag{20}$$

(20) clearly shows that phase errors can cause mainlobe misorientation when $\Phi_T > \zeta$. To keep the mainlobe orientation at 0°, Φ_T should be no larger than ζ. By examining the numerator elements of ζ, we can find four robustness boosting factors. Firstly, M suggests that more number of microphones can boost the robustness of first-order LDMAs. Secondly, $\omega = 2\pi f$ suggests that first-order LDMAs are more robust at higher frequency bands. Thirdly, $\tau_0 = \delta/c_0$ suggests that larger microphone spacing helps with array robustness. Lastly, $a_{1,0}$ suggests the choice of beampattern coefficients plays a role in array robustness.

Based on (11), we can further derive the WNG formula with the presence of phase errors as

$$\text{WNG}[\mathbf{h}(\omega)] = \frac{|\mathcal{B}_{M,1}^{(p)}(\theta_d)|^2}{[\mathbf{h}(\omega)]^H\mathbf{h}(\omega)}. \tag{21}$$

Apply $\Phi_T \leq \zeta$ and $\theta_d = 0^\circ$ into (21), we get

$$\text{WNG}[\mathbf{h}(\omega)] \geq \frac{a_{1,1}^2 + a_{1,0}^2\psi_1(\omega)^2}{[\mathbf{h}(\omega)]^H\mathbf{h}(\omega)}. \tag{22}$$

By letting $\psi_1(\omega) = \epsilon_p$, we can get

$$\text{WNG}_T[\mathbf{h}(\omega)] = \frac{a_{1,1}^2 + a_{1,0}^2\epsilon_p^2}{[\mathbf{h}(\omega)]^H\mathbf{h}(\omega)}. \tag{23}$$

(23) offers a practical robustness threshold of first-order LDMAs. As long as the tested WNG value of a first-order LDMA is higher than $\text{WNG}_T[\mathbf{h}(\omega)]$, it will not suffer from mainlobe misorientation with the presence of ϵ_p-bounded phase errors.

3.2. Analysis of Unbounded Phase Errors

For unbounded phase errors, we can rewrite (14) without approximating phase delay terms, which becomes

$$\begin{aligned}&\mathcal{B}_{M,1}^{(p')}(\theta)\\ &\approx a_{1,0}e^{\jmath\psi_1} + \frac{3a_{1,1}}{\omega\tau_0(2M-1)}\Big[\frac{2\sum_{m=2}^{M}(m-1)e^{\jmath\psi_m}}{\jmath M(M-1)} + \jmath e^{\jmath\psi_1}\Big]\\ &+ \frac{6a_{1,1}\sum_{m=2}^{M}(m-1)^2 e^{\jmath\psi_m}\cos\theta}{M(M-1)(2M-1)}\end{aligned} \tag{24}$$

By combining (17) and (24), we get the derivative of $|\mathcal{B}_{M,1}^{(p')}(\theta)|^2$ with respect to θ:

$$\frac{d|\mathcal{B}_{M,1}^{(p')}(\theta)|^2}{d\theta} = \frac{-12a_{1,1}\sum_{m=2}^{M}(m-1)^2 e^{\jmath\psi_m}}{M(M-1)(2M-1)}\mathcal{B}_{M,1}^{(p')}(\theta)\sin\theta. \tag{25}$$

Similar to the analysis of (17), from (25) we observe that we only need to consider $|\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2 \geq |\mathcal{B}_{M,1}^{(p')}(180^\circ)|^2$ to avoid mainlobe misorientation. The difference $\mathcal{D} = |\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2 - |\mathcal{B}_{M,1}^{(p')}(180^\circ)|^2$ can be derived from (24):

$$\begin{aligned}\mathcal{D} &= \frac{24a_{1,0}a_{1,1}\sum_{m=2}^{M}(m-1)^2\cos(\psi_1-\psi_m)}{M(M-1)(2M-1)} -\\ &\frac{72a_{1,1}^2\sum_{m=2}^{M}(m-1)^2\sin(\psi_1-\psi_m)}{\omega\tau_0 M(M-1)(2M-1)^2} -\\ &\frac{144a_{1,1}^2\sum_{m_1=2}^{M}\sum_{m_2=2}^{M}(m_1-1)(m_2-1)^2\sin(\psi_{m_2}-\psi_{m_1})}{\omega\tau_0 M^2(M-1)^2(2M-1)^2}.\end{aligned} \tag{26}$$

(26) shows that unbounded phase errors can also cause mainlobe misorientation if $\mathcal{D} < 0$.

Denote $|\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2_{\text{max}}$ as the maximal value of $|\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2$. Define $\psi_T = \psi_{m'} - \psi_1$ to be the theoretical phase difference that yields $|\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2_{\text{max}}$, where $m' = 2, 3, \cdots, M$. By examination, we derive

$$\psi_T = n\pi - \tan^{-1}\left[\frac{a_{1,0}\omega\tau_0(2M-1)}{3a_{1,1}}\right]. \tag{27}$$

It follows that $|\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2_{\text{max}}$ is deduced as

$$\begin{aligned}&|\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2_{\text{max}}\\ &= a_{1,0}^2 + \frac{18a_{1,1}^2}{\omega^2\tau_0^2(2M-1)^2} + \frac{6a_{1,0}a_{1,1}\sin(\psi_T)}{\omega\tau_0(2M-1)}\\ &- \frac{18a_{1,1}^2\cos(\psi_T)}{\omega^2\tau_0^2(2M-1)^2} + a_{1,1}^2 + \frac{1}{2}\mathcal{D} \qquad (28)\\ &\geq a_{1,0}^2 + \frac{18a_{1,1}^2(1-\cos(\psi_T))}{\omega^2\tau_0^2(2M-1)^2} + \frac{6a_{1,0}a_{1,1}\sin(\psi_T)}{\omega\tau_0(2M-1)}\\ &+ a_{1,1}^2 = |\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2_{T'}, \qquad (29)\end{aligned}$$

where (29) is obtained by applying $\mathcal{D} \geq 0$ into (28). Consequently, the maximal WNG threshold for unbounded phase errors is

$$\text{WNG}_{T'}[\mathbf{h}(\omega)] = \frac{|\mathcal{B}_{M,1}^{(p')}(0^\circ)|^2_{T'}}{[\mathbf{h}(\omega)]^H\mathbf{h}(\omega)}. \tag{30}$$

With unbounded phase errors, (29) and (30) indicate that more microphones enhance the robustness of LDMAs.

4. EXPERIMENTAL RESULTS

Without loss of generality, we use the first microphone in LDMAs as the reference for phase errors. Therefore, ψ_1 is 0 throughout our experiments. We configure a baseline setup of first-order LDMAs, which sets $f = 1$ kHz, $M = 3$ and $\delta = 0.01$ m.

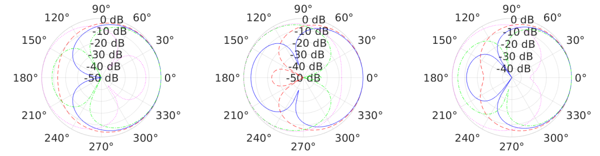
The beampatterns with and without bounded phase errors for the baseline settings are demonstrated in Fig. 1. The blue solid lines represent beampatterns without phase errors. Φ_T values adopted are listed in Table 1, where 'C', 'H' and 'S' represent cardioid, hypercardioid and supercardioid respectively. In Fig. 2, vertical lines are ζ values computed by using (20). Φ_T values on the left of corresponding ζ lines do not cause mainlobe misorientation. Combine Fig. 1 and 2. We can see that whenever $\Phi_T < \zeta$, the mainlobe orientation stays at $0°$.

From Fig. 2, we can also observe that cardioid has the largest tolerance threshold ζ and its WNG values are always larger than the other two beampatterns given the same Φ_T values. Negative Φ_T values indicate that the phase differences between microphones have not disturbed the original order of microphones for receiving signals. In other words, the effect of having negative Φ_T values is equivalent to enlarging the inter-microphone spacing. Therefore, the WNG values increase when the Φ_T values become more negative. It is interesting to note that the DF values of cardioid and supercardioid increase a bit for small positive Φ_T values. This can be corroborated by the beampattern shapes in Fig. 1(i) and 1(iii). We can see that the beampatterns drawn in green dash-dotted lines have narrower mainlobes than the beampatterns drawn in blue solid lines. The DF rebounce in Fig. 2(ii) can be also explained by the corresponding beampatterns in Fig. 1(ii). The beampattern drawn in a green dash-dotted line has more directivity than the beampattern drawn in a pink dotted line at $180°$. This suggests the opposite is true at $0°$.

Fig. 3 shows that unbounded phase errors can also cause mainlobe misorientation. Deploying more microphones can increase the WNG threshold $\text{WNG}_{T'}$, which enhances the robustness of first-order LDMAs. Notably, the $\Delta\psi$ is 0.154 for cardioid according to [14], which can be refuted by ψ_r and ψ_g. This demonstrates mainlobe misorientation can be avoided even when some phase errors are larger than $\Delta\psi$.

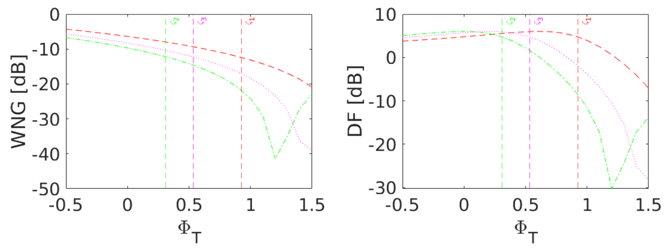
Table 1. Φ_T values used in Fig. 1.

	Red Dashed Line	Green Dash-dotted Line	Pink Dotted Line
C	-0.5	0.5	1.5
H	-0.5	1.2	1.5
S	-0.5	0.3	1.5



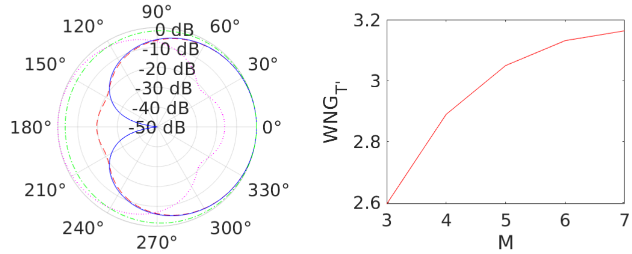
(i) Cardioid (ii) Hypercardioid (iii) Supercardioid

Fig. 1. Beampatterns of baseline settings with bounded phase errors. Φ_T values used in the figure can be found in Table 1.



(i) WNG vs Φ_T (ii) DF vs Φ_T

Fig. 2. WNG and DF plots of baseline settings with bounded phase errors. Red dashed line: cardioid. Green dash-dotted line: hypercardioid. Pink dotted line: supercardioid.



(i) Beampatterns (ii) $\text{WNG}_{T'}$ vs M

Fig. 3. Cardioid beampatterns of baseline settings with unbounded phase errors. $\psi_r = [0\ 12\ 0.2]$, $\psi_g = [0\ -10\ 0.5]$, $\psi_f = [0\ -0.1\ -7]$.

5. CONCLUSION

This paper examined the effects of phase errors on first-order LDMAs. We derived the tolerance threshold of bounded phase errors and WNG thresholds of both bounded and unbounded phase errors. In practice, the mainlobe misorientation due to bounded phase errors can be detected by using the tolerance threshold. When the tolerance threshold is exceeded, we know that mainlobe misorientation happens. Our WNG thresholds of both bounded and unbounded phase errors can serve as design criteria for the robustness of first-order LDMAs. If the WNG values are above the WNG thresholds, mainlobe misorientation of first-order LDMAs is guaranteed not to happen.

6. REFERENCES

[1] Gary W Elko, "Microphone array systems for hands-free telecommunication," *Speech communication*, vol. 20, no. 3-4, pp. 229–240, 1996.

[2] W Herbordt, T Horiuchi, M Fujimoto, T Jitsuhiro, and S Nakamura, "Hands-free speech recognition and communication on pdas using microphone array technology," in *IEEE Workshop on Automatic Speech Recognition and Understanding, 2005.* IEEE, 2005, pp. 302–307.

[3] Istvan I Papp, Zoran M Saric, and Nikola Dj Teslic, "Hands-free voice communication with tv," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 2, pp. 606–614, 2011.

[4] Kenichi Kumatani, John McDonough, and Bhiksha Raj, "Microphone array processing for distant speech recognition: From close-talking microphones to far-field sensors," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 127–140, 2012.

[5] Wenyu Jin, Mohammad J Taghizadeh, Kainan Chen, and Wei Xiao, "Multi-channel noise reduction for hands-free voice communication on mobile phones," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 506–510.

[6] Jacob Benesty and Jingdong Chen, *Study and design of differential microphone arrays*, vol. 6, Springer Science & Business Media, 2012.

[7] Jacob Benesty, Jingdong Chen, and Chao Pan, *Fundamentals of differential beamforming*, Springer, 2016.

[8] Enzo De Sena, Hüseyin Hacihabiboglu, and Zoran Cvetkovic, "On the design and implementation of higher order differential microphones," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 1, pp. 162–174, 2011.

[9] Liheng Zhao, Jacob Benesty, and Jingdong Chen, "Design of robust differential microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, pp. 1455–1466, 2014.

[10] Simon Doclo and Marc Moonen, "Superdirective beamforming robust against microphone mismatch," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. IEEE, 2006, vol. 5, pp. V–V.

[11] Markus Buck, "Aspects of first-order differential microphone arrays in the presence of sensor imperfections," *European transactions on telecommunications*, vol. 13, no. 2, pp. 115–122, 2002.

[12] Quansheng Tu and Huawei Chen, "Theoretical lower bounds on the performance of the first-order differential microphone arrays with sensor imperfections," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2022.

[13] Quansheng Tu and Huawei Chen, "On mainlobe orientation of the first-and second-order differential microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 12, pp. 2025–2040, 2019.

[14] Zuolong Chen, Huawei Chen, and Quansheng Tu, "Sensor imperfection tolerance analysis of robust linear differential microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2915–2929, 2021.

[15] Mohamad Hasan Bahari, Alexander Bertrand, and Marc Moonen, "Blind sampling rate offset estimation for wireless acoustic sensor networks through weighted least-squares coherence drift estimation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 3, pp. 674–686, 2017.

[16] Tobias Gburrek, Joerg Schmalenstroeer, and Reinhold Haeb-Umbach, "Geometry calibration in wireless acoustic sensor networks utilizing doa and distance information," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2021, no. 1, pp. 1–17, 2021.

[17] Aleksej Chinaev, Gerald Enzner, Tobias Gburrek, and Joerg Schmalenstroeer, "Online estimation of sampling rate offsets in wireless acoustic sensor networks with packet loss," in *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 1110–1114.

[18] Tobias Gburrek, Joerg Schmalenstroeer, and Reinhold Haeb-Umbach, "On synchronization of wireless acoustic sensor networks in the presence of time-varying sampling rate offsets and speaker changes," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 916–920.

[19] Barry D Van Veen and Kevin M Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE assp magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[20] Jacob Benesty, Jingdong Chen, and Yiteng Huang, *Microphone array signal processing*, vol. 1, Springer Science & Business Media, 2008.

[21] EN Gilbert and SP Morgan, "Optimum design of directive antenna arrays subject to random variations," *Bell System Technical Journal*, vol. 34, no. 3, pp. 637–663, 1955.