Gradient Flow Source Localization in Noisy and Reverberant Environments

Shuo Li and Milutin Stanaćević
Department of Electrical and Computer Engineering
Stony Brook University
Stony Brook, NY 11794–2350
Email: shuoli,milutin@ece.sunysb.edu

***Abstract*— Gradient flow is a technique for localization of an acoustic source using miniature microphone arrays by relating temporal and spatial gradients of the impinging source signal. The performance of the gradient flow in noisy and reverberant environment is examined and quantified through simulations that incorporate additive measurement noise, directional interference signal and room acoustic model. The algorithm demonstrates robust performance for additive signal-to-noise ratio down to 5 dB and at the signal-to-interference ratio of 10 dB. In the echoic room, localization performance starts to deteriorate under moderate reverberation conditions. The experimental results from a miniature microphone array recordings in a conference-room environment verify the presented simulations.**

I. INTRODUCTION

Performance of source localization algorithms degrades in the presence of multiple noise sources and is typically limited by room reverberation. To access the localization performance under these adverse conditions is of great interest to a variety of applications in hearing enhancement, multimedia, communication and surveillance. Traditionally, sensor arrays with large inter-sensor distances are required to guarantee sufficient spatial separation across sensors and the source localization algorithms are mostly based on estimation of the time delays between source observations [1], [2]. The effect of the reverberation on the performance of these types of algorithms has been studied extensively, with the conclusion that the performance significantly deteriorates above certain level of room reverberation [3], [4].

However, many applications call for localization of acoustic sources using miniature microphone arrays, where resolving temporal difference poses a significant challenge due to the sub-wavelength distance between microphones. Inspired by optical flow for motion estimation in the field of imaging, the direction of sound propagation can be inferred directly from sensing spatial and temporal gradients of the wave signal on a sub-wavelength scale [5], [6]. This principle is also observed in biology, *e.g.,* for localizing prey transmission using differential eardrums and neuronal processing structure or reflecting high-frequency sound, and has already been implemented in biomimetic MEMS systems. Gradient flow technique lends itself into efficient implementation in low-power smart time-to-digital conversion mixed-signal VLSI system, amenable to miniature sensing and processing applications [7]. In this paper, we study the performance of the gradient flow algorithm in noisy and reverberant conditions using synthetically generated microphone array signals and compare these simulation results with the measured performance in the office-room environment.

II. GRADIENT FLOW ACOUSTIC LOCALIZATION

Gradient flow [5], [7] is a signal conditioning technique for source localization designed for arrays of very small aperture, *i.e.,* of dimensions significantly smaller than the shortest wavelength in the sources. Consider a traveling acoustic wave impinging on an array of four microphones, in the configuration of Figure 1. The 3-D direction cosines of the traveling wave $\mathbf{u}$ are implied by propagation delays τ_1 and τ_2 in the source along directions p and q in the sensor plane. Direct measurement of these delays is problematic as they require sampling in excess of the bandwidth of the signal, increasing noise floor and power requirements. However, indirect estimates of the delays are obtained, to first order, by relating spatial and temporal derivatives of the acoustic field. The signal impinging at the sensor with position coordinates p and q can be expanded about the center of the array in the power series expansion as

$$x_{pq}(t) = s(t) + \tau_{pq}\dot{s}(t) + \frac{1}{2}(\tau_{pq})^2\ddot{s}(t) + ... + n_{pq}(t) \tag{1}$$

We concentrate on the first two terms in the series expansion, which is linear in the space coordinates.

$$x_{pq}(t) \approx s(t) + \tau_{pq}\dot{s}(t) + n_{pq}(t) \tag{2}$$

We take spatial derivatives of various orders i and j along the coordinates p and q, round the origin $p = q = 0$, and apply the above series expansion model to obtain:

$$\begin{aligned} \xi_{ij}(t) &\equiv \frac{\partial^{i+j}}{\partial^i p \partial^j q} x_{pq}(t) \\ &= (\tau_1)^i(\tau_2)^j \frac{d^{i+j}}{d^{i+j}t} s(t) + \nu_{ij}(t) \end{aligned} \tag{3}$$

where ν_{ij} is the corresponding spatial derivatives of the sensor noise n_{pq} around the center. For the first-order case which $i + j = 1$, the equation above becomes:

$$\begin{aligned} \xi_{10}(t) &\approx \tau_1\dot{\xi}_{00}(t) \\ \xi_{01}(t) &\approx \tau_2\dot{\xi}_{00}(t) \end{aligned} \tag{4}$$

where ξ_{10} and ξ_{01} represent spatial gradients in p and q directions around the origin ($p = q = 0$), ξ_{00} the spatial

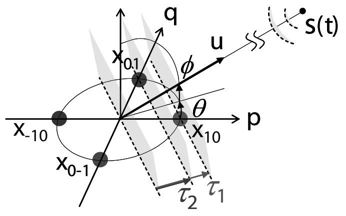

Fig. 1. Configuration of sensors for spatial gradient estimation.

common mode, and $\dot{\xi}_{00}$ its time derivative. Estimates of ξ_{00}, ξ_{10} and ξ_{01} are obtained by finite difference gradient approximation on a grid from the sensor observations $x_{-1,0}$, $x_{1,0}$, $x_{0,-1}$ and $x_{0,1}$ as:

$$\begin{aligned} \xi_{00} &\approx \tfrac{1}{4}\big(x_{-1,0} + x_{1,0} + x_{0,-1} + x_{0,1}\big) \\ \xi_{10} &\approx \tfrac{1}{2}\big(x_{1,0} - x_{-1,0}\big) \\ \xi_{01} &\approx \tfrac{1}{2}\big(x_{0,1} - x_{0,-1}\big) \end{aligned} \tag{5}$$

From least-square estimates of the time delays in (4), we directly obtain bearing estimates of azimuth angle θ and elevation angle ϕ of $\mathbf{u}$ relative to p and q as

$$\begin{aligned} \tau_1 &= \frac{1}{c}|\mathbf{r}_1|\ \cos\theta\sin\phi \\ \tau_2 &= \frac{1}{c}|\mathbf{r}_2|\ \sin\theta\sin\phi, \end{aligned} \tag{6}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are unit orthogonal vectors along p and q directions.

III. Simulated Localization Performance

The performance of gradient flow algorithm for source localization is quantified in simulated adverse acoustic conditions. We performed simulations with artificially generated microphone array signals with scaled additive measurement noise and different levels of background noise with different incidence angle of the source signal. The simulated room experiments were performed to determine the degrading effect of reverberation in the real room environment.

A. Localization Performance with Additive Measurement Noise and Directional Noise Source

We first investigate the effect of the additive measurement noise, that can originate from the sensor itself or the readout electronics. The assumption is that the mutually independent white Gaussian noise sources are scaled and added to each microphone signal to control the signal-to-noise ratio (SNR). The performance of the localization algorithm was analyzed given the results of a series of Monte Carlo simulation experiments with artificially generated source signals and additive noise. We use two different types of the source signals, one signal is a bandlimited (100Hz-4kHz) white Gaussian source signal, while the second source is the speech signal from the TIMIT database. The sampling frequency is 16 kHz, while the length of the signal is 1 s. The distance between microphones is set to 1 cm. The signal-to-noise ratio at the received sensor signals is varied from -5dB to 15dB for the incidence angle of the source

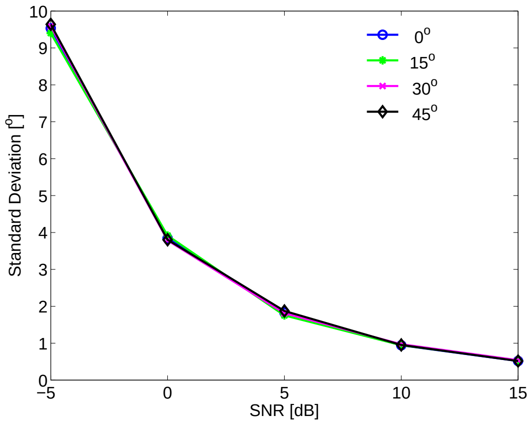

Fig. 2. Standard deviation error of the estimated incidence angle under different SNR conditions and at different angular location of the source for bandlimited white Gaussian signal.

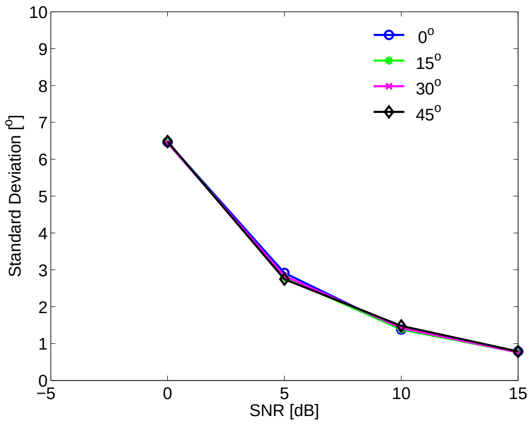

Fig. 3. Standard deviation error of the estimated incidence angle under different SNR conditions and at different angular location of the source for speech signal.

signal of 0^o, 15^o, 30^o and 45^o. The choice of angles is justified by the symmetry of the four element microphone array. The Figure 2 shows the standard deviation of localization error for different angles for band-limited white Gaussian random source, while the Figure 3 shows the standard deviation error for a speech source signal.

Despite the low aperture, the gradient flow demonstrates a robust performance with respect to acquisition noise. The performance is maintained to level of SNR as low as 5 dB, while it deteriorates significantly when the SNR drops down to 0 dB. We can also observe that the localization performance is not influenced by the incidence angle.

In the next set of simulations, we investigate the effect of a directional noise sound source on the localization performance. The source signal is a speech signal chosen from TIMIT database, while the interference source is a bandlimited white Gaussian signal. The source signal is located at incidence angle of 45^o, while the interferer is moved from the angle of 30^o to the angle of -45^o in step of 15^o, thus varying the angular distance from 15^o of separation to 90^o. The signal-

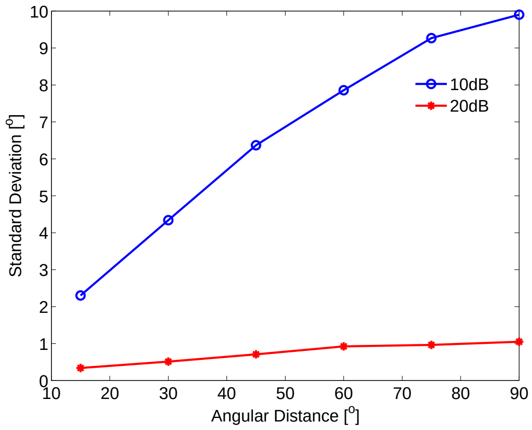

Fig. 4. Standard deviation error of the estimated incidence angle if an interfering source is present in the environment.

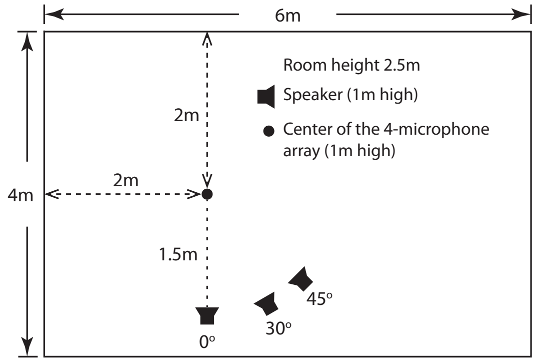

Fig. 5. Room dimension.

to-interference ratio (SIR) is set at 10dB and 20dB, with the additive measurement noise of 20dB. For the SIR of 20dB the influence of the interferer on the localization performance is minimal, while at 10dB of SIR the influence of the interferer dominates the standard deviation of the error.

B. Localization Performance under Reverberant Conditions

We want to understand how room reverberation affects the localization performance of the gradient flow technique. In the room environment, in addition to the direct-path signal received by the microphone array, the observed signals comprise multiple delayed and attenuated replicas of the source signal due to reflections from the room boundaries as walls, ceiling, and floor. The reflections also originate from different objects present in the room. The observer perceives the echo as radiating from a point past the wall from which it was reflected.

To quantify the localization performance under different reverberant conditions, we construct an artificial room using virtual source mapping model to generate observations at the microphone array in an echoic room environment. The image method [8] is used to generate the imaginary sources, with their location and the incidence angle on the microphone array. The reflected sound is propagated from the direction of the sound source of the mirror image room, and the image of the source in the mirror room is referred to as a virtual source. If we get the location and number of reflections of the virtual sources and the reflection coefficient, we can get an impulse response of the signal source in an echoic environment. In our experiment, corresponding room impulse responses is generated based on the calculated reflection coefficient according to the room dimensions and reverberation time. The signal received at the microphone x_{10} is represented as

$$x_{10}(t) = \sum_{i=1}^{N} a_i s(t - \tau^i) + n_{10}(t). \quad (7)$$

where a_i and τ^i are attenuation and time delay associated with the virtual source i and N is the number of virtual sources. We assume that the source location is fixed, so that the model is time-invariant. The dimensions of the room, with the location of the microphone array and the source signal, are illustrated in Figure 5. The distance between the array and source was kept at 1.5 m, while we investigate three different incidence source angles, 0^o, 30^o and 45^o. The additive measurement noise was mutually independent, white, Gaussian random signal with power varied so that the signal-to-noise ratio was varied from 0dB to 20dB. The reflection coefficient is assumed to be frequency-independent and the same for all walls, ceiling and floor. For a given value of reverberation time, T_{60}, the reflection coefficient is computed using Eyring's formula:

$$r = \exp\left(-\frac{13.82}{(\frac{1}{\mathrm{L_x}} + \frac{1}{\mathrm{L_y}} + \frac{1}{\mathrm{L_z}})\mathrm{cT_{60}}}\right). \quad (8)$$

Figure 6 and Figure 7 show the dependance of accuracy in the estimated angle of incidence, as the percentage of the total number of absolute errors smaller than 5^o and standard deviation, as a function of the reverberation time for different incidence angles and different level of additive measurement noise. We can conclude that the gradient flow demonstrates robust performance under mild reverberation. Under moderate reverberation, the gradient flow performs well if the additive noise is on the order of 10 dB. As expected, the impact of the reverberation is strongest for the incidence angle of 45^o.

IV. Experimental Results

To verify the simulation results and to demonstrate the performance of the gradient flow algorithm, the planar array of four omnidirectional hearing aid miniature microphones was used for localization experiments in the room environment. A single acoustic source was presented through a loudspeaker positioned at 1 m distance from the array in a conference room with reverberation time of 180 ms, while the microphone array was mounted on a rotating platform. The localization experiment was repeated for two different source signals, a broadband bandlimited (100-1000Hz) Gaussian signal and speech signal. The corresponding signal-to-noise(SNR) ratio was around 35 dB. The 10 estimates of the incidence angle were obtained for 1 s long source signals. The mean and variance of estimated incidence angles for Gaussian and speech signal are shown in Figure 8 and Figure 9.

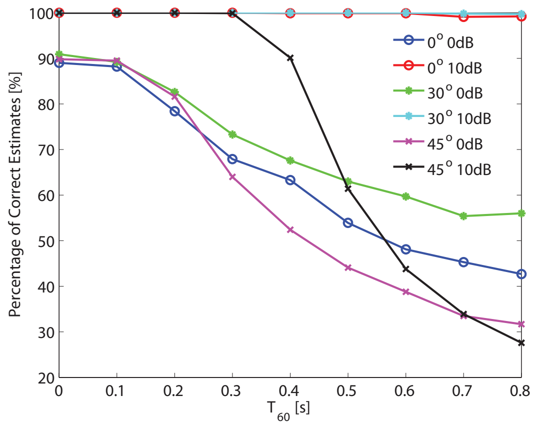

Fig. 6. Percentage of correctness in the estimated angle of incidence as a function of the reverberation time for different incidence angle and different level of additive measurement noise.

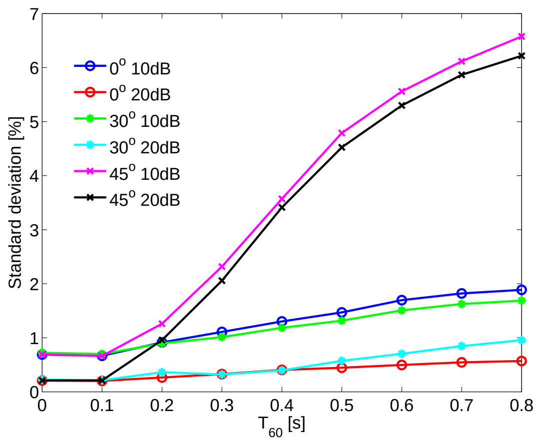

Fig. 7. Standard deviation of the estimated angle of incidence as a function of the reverberation time for different incidence angle and different level of additive measurement noise.

V. Conclusion

The gradient flow acoustic localization is presented and evaluated in adverse noisy and reverberant environments. The simulation results demonstrate robust performance with mild to moderate reverberations with additive noise levels down to 10dB. As the gradient flow technique in combination with independent component analysis provides a unique approach to the task of blind acoustic source separation, the separation performance will be investigated and quantified in the similar environments using the models developed in the presented work.

VI. Acknowledgment

This work was supported by NSF CAREER Award 0846265.

References

[1] Y. Huang, J. Benesty, G.W. Elko and R.M. Mersereau, "Real-Time Passive Source Localization: A Practical Linear-Correction Least-Squares Approach", *IEEE Transactions on Speech and Audio Procession*, vol. **9** (8), pp. 943-956, 2001.

[2] D. Halupka, N.J. Mathai, P. Aarabi and A. Sheikholeslami, "Robust Sound localization in 0.18 μm CMOS", *IEEE Transactions on Signal Processing*, vol. **53** (6), pp. 2243-2250, 2005.

[3] B. Champagne, S. Bedard and A. Stephenne, "Performance of Time-Delay Estimation in the Presence of Room Reverberation", *IEEE Transactions on Speech and Audio Procession*, vol. **4** (2), pp. 148-152, 1996.

[4] T. Gustafsson, B.D. Rao and M. Trivedi, "Source Localization in Reverberant Environments: Modeling and Statistical Analysis", *IEEE Transactions on Speech and Audio Procession*, vol. **11** (6), pp. 791-803, 2003.

[5] G. Cauwenberghs, M. Stanaćević, and G. Zweig, "Blind Broadband Source Localization and Separation in Miniature Sensor Arrays," *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'2001),* Sydney, Australia, May 6-9, 2001.

[6] M. Stanaćević, G. Cauwenberghs and G. Zweig, "Gradient Flow Adaptive Beamforming and Signal Separation in a Miniature Microphone Array," *Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing* (ICASSP'2002), Orlando FL, May 13-17, 2002.

[7] M. Stanaćević and G. Cauwenberghs, "Micropower Gradient Flow VLSI Acoustic Localizer", *IEEE Transactions on Circuits and Systems I : Regular Papers*, vol. **52** (10), pp. 2148-2157, 2005.

[8] J.B. Allen and D.A. Berkley, "Image method for efficiently simulating small-room acoustics", *J. Acoust. Soc. Amer.*, vol. **65**, pp. 943950, Apr. 1979.

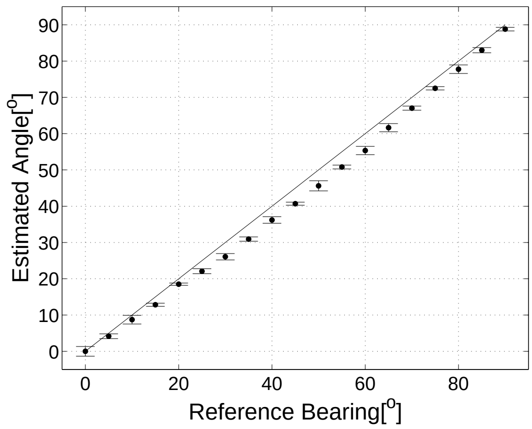

Fig. 8. Localization of bandlimited (100-1000Hz) Gaussian source in room environment.

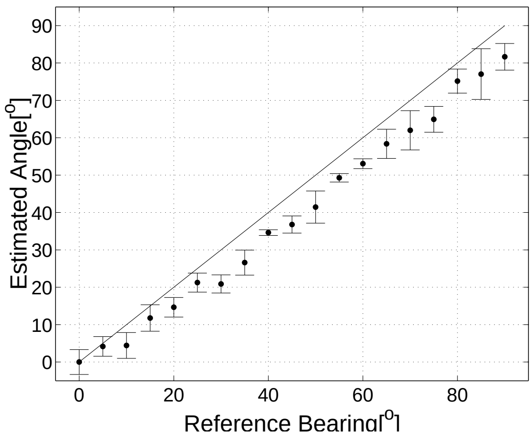

Fig. 9. Localization of speech source signal in room environment.