

# Micropower Gradient Flow Acoustic Localizer

Milutin Stanačević, *Student Member, IEEE*, and Gert Cauwenberghs, *Senior Member, IEEE*

**Abstract**—A micropower mixed-signal system-on-chip for three-dimensional localization of a broad-band acoustic source is presented. Direction cosines of the source are obtained by relating spatial and temporal differentials in the acoustic traveling wave field acquired across four coplanar microphones at subwavelength spacing. Correlated double sampling and least-squares adaptive cancellation of common-mode leakthrough in the switched-capacitor analog differentials boost localization accuracy at very low aperture. A second stage of mixed-signal least-squares adaptation directly produces digital estimates of the direction cosines. The 3 mm × 3 mm chip in 0.5- $\mu\text{m}$  CMOS technology quantizes signal delays with 250-ns resolution at 16-kHz sampling rate, and dissipates 54  $\mu\text{W}$  power from a 3-V supply. Field tests of the processor with acoustic enclosure demonstrated its utility and endurance in tracking ground and airborne vehicles. Applications include acoustic surveillance, interactive multimedia, and intelligent hearing aids.

**Index Terms**—Array signal processing, correlated double sampling (CDS), least-mean-squares (LMS) adaptation, micropower techniques, source localization, switched-capacitor (SC) circuits.

## I. INTRODUCTION

**P**RECISE and robust localization and tracking of acoustic sources is of interest to a variety of applications in surveillance, multimedia, and hearing enhancement. Miniaturization of microphone arrays integrated with acoustic processing further increases the utility of these systems, but poses challenges to attain accurate localization performance due to shrinking aperture [1]. For surveillance, acoustic emissions from ground vehicles provide an easily detected signature that can be used for unobtrusive and passive tracking. In multimedia, localization of a speaker can be used for steering in automatic teleconferencing, hands-free portable devices and robot navigation. In hearing aids and speech recognition systems, the presence of multiple signal and noise sources in the acoustic scene limits speech intelligibility and necessitates localization and separation of multiple acoustic sources. There have been a number of VLSI systems for acoustic direction finding reported in the literature [2]–[8].

An integrated miniature sensor array with localization and communication capability could be contained as a low-cost, low-power small autonomous node in network configuration distributed over a wide area [9]–[11]. This leads to higher localization performance in distributed sensing environments bypassing the need for excessive data transfer and fine-grain

Manuscript received January 24, 2005. This work was supported by the Office of Naval Research under Grant N00014-99-1-0612, Grant ONR/DARPA N00014-00-1-0838 and Grant NSF/DIA IIS-0434161. The chip was fabricated through the MOSIS service. This paper was recommended by Associate Editor G. Palumbo.

The authors are with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: miki@jhu.edu).  
Digital Object Identifier 10.1109/TCSI.2005.853356

time synchronization among nodes, with low communication bandwidth and low complexity. Additional improvement can also be achieved by fusion with other data modalities, like video [12].

Conventionally, large sensor arrays are used for source localization to warrant sufficient spatial diversity across sensors to resolve time delays between source observations. It is well known that the precision of delay-based bearing estimation degrades with shrinking dimensions (aperture) of the sensor array [1], [13]. Time-difference of arrival estimation techniques based on cross-correlation of the signals [14] require high oversampling ratios for estimating small time delays [15]. Gradient flow [16], [17] is a signal conditioning technique for arrays of very small dimensions, which converts time delays between signal observations into relative amplitudes of the time-differentiated signal, by observing gradients (spatial differences). Improved differential sensitivity of gradient sensing allows to shrink the aperture of the sensor array without degrading signal-to-noise ratio. Interestingly, some insects are capable of remarkable auditory localization at dimensions a small fraction of the wavelength, owing to differential processing of sound pressure through inter-tympanal mechanical coupling [18] or inter-aural coupled neural circuits [19]. Besides its use in bearing estimation, gradient flow provides an efficient signal representation as a front-end for blind source separation using independent component analysis [20].

The paper is organized as follows. Section II reviews the gradient flow approach to localizing a single source in the free field. Section III presents a mixed-signal architecture implementing gradient flow, including computation of spatial and temporal gradients using finite differences on an array of four sensors in the plane and least-squares adaptation for common-mode rejection and bearing estimation. Experimental results on a fabricated prototype [21], in addition to characterization of localization performance and field tests for ground vehicle tracking are presented in Section IV. Concluding arguments are given in Section V.

## II. GRADIENT FLOW LOCALIZATION

A traveling wave emitted by a source is observed over a distribution of sensors in space, which here we consider to be discrete but which could be continuous. We define  $\tau(\mathbf{r})$  as the time advance between the wavefront at point  $\mathbf{r}$  and the wavefront at the center of the array, i.e., the propagation time  $\tau(\mathbf{r})$  is referenced to the center of the array.

In the *far field*, the wavefront delay  $\tau(\mathbf{r})$  is approximately linear in the projection of  $\mathbf{r}$  on the unit vector  $\mathbf{u}$  pointing toward the source

$$\tau(\mathbf{r}) \approx \frac{1}{c} \mathbf{r} \cdot \mathbf{u} \quad (1)$$

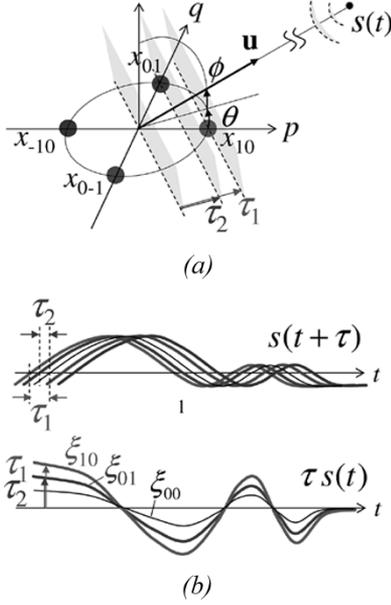


Fig. 1. Illustration of gradient flow principle. (a) Sensor geometry and source direction coordinates. (b) *Top*: sensor observations *Bottom*: spatial gradients.

where  $c$  is the speed of wave propagation. In particular we consider a two-dimensional array of sensors, with position coordinates  $p$  and  $q$  so that  $\mathbf{r}_{pq} = p\mathbf{r}_1 + q\mathbf{r}_2$  with orthogonal vectors  $\mathbf{r}_1$  and  $\mathbf{r}_2$  in the sensor plane. In the far-field approximation (1), the sensor observations of the source are advanced in time by  $\tau_{pq} = p\tau_1 + q\tau_2$ , where

$$\begin{aligned}\tau_1 &= \frac{1}{c} \mathbf{r}_1 \cdot \mathbf{u} \\ \tau_2 &= \frac{1}{c} \mathbf{r}_2 \cdot \mathbf{u}\end{aligned}\quad (2)$$

are the interaural time differences (ITD) of source between adjacent sensors on the grid along the  $p$  and  $q$  place coordinates, respectively. These ITDs (delays)  $\tau_1$  and  $\tau_2$  thus represent the *direction cosines* (or *angle coordinates*) of the source relative to the array, scaled by the speed of wave propagation and the unit dimensions of the array. The direction cosines can be decomposed in azimuth angle  $\theta$  and elevation angle<sup>1</sup>  $\phi$  as indicated in Fig. 1(a)

$$\begin{aligned}\tau_1 &= \frac{1}{c} |\mathbf{r}_1| \cos \theta \sin \phi \\ \tau_2 &= \frac{1}{c} |\mathbf{r}_2| \sin \theta \sin \phi.\end{aligned}\quad (3)$$

The signal observed at sensor with position coordinates  $p$  and  $q$  can be expressed as

$$x_{pq}(t) = s(t + \tau_{pq}) + n_{pq}(t) \quad (4)$$

<sup>1</sup>To remove the indeterminacy in the sign of the elevation angle  $\phi$ , we assume that the sources impinge on top, not on bottom, of the array. This is a reasonable assumption for an integrated MEMS or VLSI array since the substrate masks any source impinging from beneath.

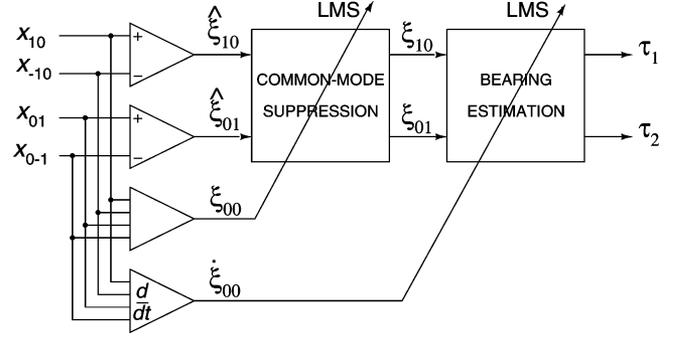


Fig. 2. System block diagram.

where  $n_{pq}(t)$  represent additive noise in the sensor observations. Gradient flow evaluates spatial gradients of the field  $x_{pq}$  along the  $p$  and  $q$  position coordinates, around the origin  $p = q = 0$

$$\begin{aligned}\xi_{ij}(t) &\equiv \left. \frac{\partial^{i+j}}{\partial^i p \partial^j q} x_{pq}(t) \right|_{p=q=0} \\ &= (\tau_1)^i (\tau_2)^j \frac{d^{i+j}}{dt^{i+j}} s(t) + \nu_{ij}(t)\end{aligned}\quad (5)$$

where  $\nu_{ij}$  are the corresponding spatial derivatives of the sensor noise  $n_{pq}$  around the center. Taking spatial derivatives  $\xi_{ij}$  of order  $i+j \leq k$ , and differentiating  $\xi_{ij}$  to order  $k-(i+j)$  in time yields a number of different linear observations in the  $k$ th-order time derivatives of the signals  $s$ . The advantage of this technique is that it effectively reduces the problem of directly estimating time delays to linear regression of the derivative signals [17].

Here, we consider the first-order case  $k = 1$

$$\begin{aligned}\xi_{00}(t) &= s(t) + \nu_{00}(t) \\ \xi_{10}(t) &= \tau_1 \dot{s}(t) + \nu_{10}(t) \\ \xi_{01}(t) &= \tau_2 \dot{s}(t) + \nu_{01}(t).\end{aligned}\quad (6)$$

As illustrated in Fig. 1(b), the spatial gradient observations reduce to amplitude scaling of the time derivative  $\dot{s}$  of the signal. The problem of localization then reduces to standard linear regression of the observed spatial and temporal derivatives

$$\begin{aligned}\xi_{10}(t) &\approx \tau_1 \dot{\xi}_{00}(t) \\ \xi_{01}(t) &\approx \tau_2 \dot{\xi}_{00}(t)\end{aligned}\quad (7)$$

in the unknown ITDs  $\tau_1$  and  $\tau_2$ .

An interesting observation from (7) is that the estimate of azimuth angle  $\theta$  according to (3) is independent of the speed of sound as it involves spatial gradients only; estimation of the elevation angle  $\phi$  on the other hand requires knowledge of the speed of sound in relating spatial and temporal derivatives.

### III. CHIP ARCHITECTURE

The system block diagram implementing gradient flow for bearing estimation is shown in Fig. 2. Spatial gradients are approximated by evaluating finite differences over the four sensors on the planar grid shown in Fig. 1(a). Accurate bearing estimation assumes accurate sensing of the gradients. Differential amplification is performed in the analog domain, as a low-power alternative to high-resolution analog-to-digital conversion and subsequent digital signal processing. Two stages of

mixed-signal adaptation compensate for common mode errors in the differential amplification, and produce digital estimates of ITDs  $\tau_1$  and  $\tau_2$  from the spatial and temporal differentials.

Estimation of the gradients is implemented using sampled-data switched-capacitor (SC) circuits. The advantage of this realization is application of correlated double sampling (CDS) that significantly reduces common-mode offsets and  $1/f$  noise [22]. The spatial gradients are computed in fully differential mode, to provide increased clock and supply feedthrough rejection. A cascoded nMOS inverter, described in Section III-D, implements a simple high-gain amplifier in SC circuits, supporting high density of integration, and high energetic efficiency.

### A. Spatial and Temporal Derivative Sensing

Gradient flow requires computation of first-order spatial and temporal derivatives of the signal impinging on the sensor array. Estimates of  $\xi_{00}$ ,  $\xi_{10}$  and  $\xi_{01}$  are obtained by finite-difference gradient approximation on a grid, using the planar array of four sensors shown in Fig. 1(a) with observations  $x_{-1,0}$ ,  $x_{1,0}$ ,  $x_{0,-1}$  and  $x_{0,1}$

$$\begin{aligned}\xi_{00} &\approx \frac{1}{4}(x_{-1,0} + x_{1,0} + x_{0,-1} + x_{0,1}) \\ \xi_{10} &\approx \frac{1}{2}(x_{1,0} - x_{-1,0}) \\ \xi_{01} &\approx \frac{1}{2}(x_{0,1} - x_{0,-1}).\end{aligned}\quad (8)$$

One significant advantage of gradient flow localization is that it relaxes requirements on sampling rate, governed by signal bandwidth (subkilohertz for acoustic surveillance) as opposed to the desired resolution of ITD estimation (microseconds). However, it is crucial to ensure that the signals be properly bandlimited, and that all signal components including spatial and temporal gradients be synchronously sampled. The timing of common-mode and gradient signal acquisition is discussed next.

1) *Common-Mode Component  $\xi_{00}$  and Temporal Derivative  $\dot{\xi}_{00}$* : The common-mode component is decomposed in differential form  $\xi_{00}[n] = \xi_{00}^+[n] - \xi_{00}^-[n]$  with

$$\begin{aligned}\xi_{00}^+[n] &= \frac{1}{8} \left( x_{10} \left[ n - \frac{1}{2} \right] + x_{-10} \left[ n - \frac{1}{2} \right] \right. \\ &\quad \left. + x_{01} \left[ n - \frac{1}{2} \right] + x_{0-1} \left[ n - \frac{1}{2} \right] \right) \\ \xi_{00}^-[n] &= -\frac{1}{8} (x_{10}[n] + x_{-10}[n] + x_{01}[n] + x_{0-1}[n]).\end{aligned}\quad (9)$$

The timing of common-mode acquisition is illustrated in Fig. 3(a). The contribution  $\xi_{00}^+$  to  $\xi_{00}$  represents the estimate of the average signal at time instance  $nT - T/2$ , while the contribution  $\xi_{00}^-$  represents the inverted estimate at time instance  $nT$ . The difference between both contributions hence produces an unbiased estimate of  $\xi_{00}$  centered at time  $nT - T/4$  (as average of the estimates at times  $nT - T/2$  and  $nT$ ).

SC realization of the common-mode estimation is given in Fig. 3(b). In the sampling phase  $\phi_1$ , input signals are sampled on the capacitors  $C_1$  in the branch computing  $\xi_{00}^+$ , while the zero-level reference voltage  $V_{\text{ref}}$  (set to  $V_{\text{dd}}/2$ ) is sampled in the branch computing  $\xi_{00}^-$ . The feedback capacitors  $C_2$  are

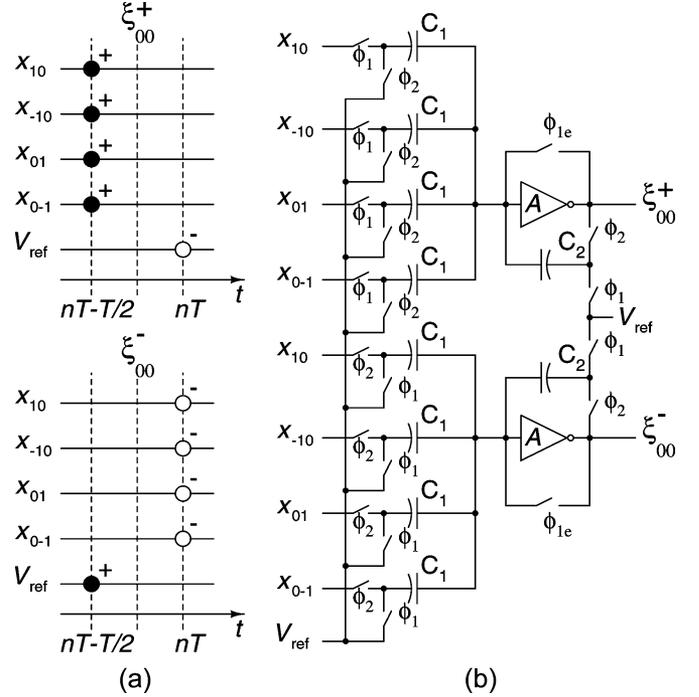


Fig. 3. (a) Estimation of spatial common-mode signal  $\xi_{00}$ . (b) SC realization.

precharged to zero-level reference voltage and the inverters are reset. In the computation phase  $\phi_2$ , the common-mode output signal is established by feedback of  $C_2$  with the order of switches in the input stage reversed, connecting sampling capacitors  $C_1$  to zero-level reference in the positive branch and to the input voltages in the negative branch. The inter-poly capacitors are sized with  $C_1$  nominally 200 fF and  $C_2$  400 fF, providing (differentially) a gain of 4 in the common-mode estimation. Capacitor sizing is governed by tradeoff between noise, slew-rate and power consumption.

The clocks  $\phi_1$  and  $\phi_2$  are nonoverlapping, and  $\phi_{1e}$  replicates  $\phi_1$  with its falling edge slightly preceding the falling edge of  $\phi_1$ . All the switches are minimum-size complementary transmission gate field-effect transistors (FETs), except the switches controlled by  $\phi_{1e}$ , which are  $n$ -channel FETs.

The estimate  $\xi_{00}$  is used for suppressing leakage of the common-mode component in the spatial gradient signals in Section III-B. Its temporal derivative  $\dot{\xi}_{00}$  is estimated independently. The estimate  $\dot{\xi}_{00}$ , centered at same time instance  $nT - T/4$ , is computed differentially in a manner similar to  $\xi_{00}$ , as illustrated in Fig. 4(a)

$$\begin{aligned}\dot{\xi}_{00}^+[n] &= 0 \\ \dot{\xi}_{00}^-[n] &= -\frac{1}{8} (x_{10}[n] + x_{-10}[n] + x_{01}[n] + x_{0-1}[n]) \\ &\quad + \frac{1}{8} \left( x_{10} \left[ n - \frac{1}{2} \right] + x_{-10} \left[ n - \frac{1}{2} \right] \right. \\ &\quad \left. + x_{01} \left[ n - \frac{1}{2} \right] + x_{0-1} \left[ n - \frac{1}{2} \right] \right).\end{aligned}\quad (10)$$

$\dot{\xi}_{00}$  is obtained as the difference of common-mode signals at time instances  $nT - T/2$  and  $nT$  and approximately equals

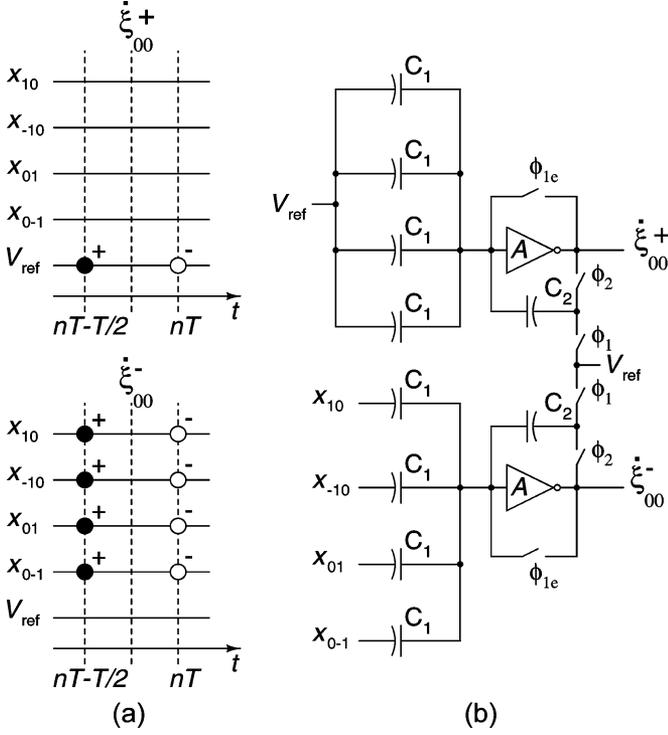


Fig. 4 (a) Estimation of temporal derivative common-mode signal  $\dot{\xi}_{00}$ . (b) SC realization.

the time derivative of the common-mode signal at  $nT - T/4$  scaled by  $T/2$ . The SC realization is shown in Fig. 4(b). The sensor signals are differenced on sampling capacitors  $C_1$  and then amplified in computing phase  $\phi_2$  on feedback capacitors  $C_2$ . Capacitor sizing  $C_1$  is 400 fF and  $C_2$  is 200 fF, providing a gain of 16 to  $\dot{\xi}_{00}$ .

2) *Gradient Components  $\xi_{10}$  and  $\xi_{01}$* : The small aperture of the sensor array poses a design challenge in resolving small signal spatial gradients amidst a large common-mode signal pedestal. Differential low-noise amplification eliminates the common-mode component and boosts differential sensitivity. The first-order spatial gradient in  $p$  direction  $\xi_{10}$ , also centered at time  $nT - T/4$ , is computed as average of  $\xi_{10}$  estimates at time instances  $nT - T/2$  and  $nT$ . For amenable implementation, contributions to  $\xi_{10}^+$  and  $\xi_{10}^-$  are distributed in the following manner:

$$\begin{aligned}\xi_{10}^+[n] &= \frac{1}{4} \left( x_{10} \left[ n - \frac{1}{2} \right] - x_{-10}[n] \right) \\ \xi_{10}^-[n] &= \frac{1}{4} \left( x_{-10} \left[ n - \frac{1}{2} \right] - x_{10}[n] \right)\end{aligned}\quad (11)$$

as illustrated in Fig. 5(a). The SC realization shown in Fig. 5(b) includes provisions for common-mode suppression, described in Section III-B. In the sampling phase  $\phi_1$ , input signal  $x_{10}$  is sampled on capacitor  $C_1$  in the branch computing  $\xi_{10}^+$ , while in the branch computing  $\xi_{10}^-$  signal  $x_{-10}$  is sampled. In the computation phase  $\phi_2$ , spatial gradient signal is computed by exchanging the inputs selected to the sampling capacitors. The size of capacitor  $C_1$  is 1 pF and capacitor  $C_2$  is 200 fF, providing gain of 20 to first-order spatial gradient. The first-order spatial gradient in the  $q$  direction,  $\xi_{01}$ , is computed in identical fashion.

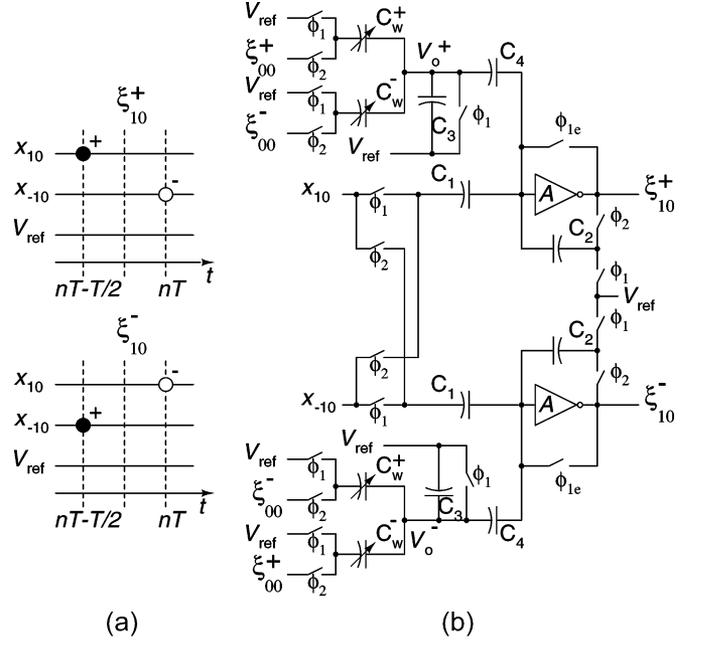


Fig. 5. (a) Estimation of first-order spatial gradient along  $p$  direction,  $\xi_{10}$ . (b) SC differential realization.

### B. Common-Mode Suppression

Mismatch across the sensor array introduces errors in estimation of the spatial gradients which translates to errors in bearing estimates. To mitigate this effect, common mode correction of the gain mismatch errors in the sensors is applied to the estimated spatial gradients, prior to bearing estimation. For  $\xi_{10}$ , gain mismatch can be represented to first order as

$$\begin{aligned}\hat{\xi}_{10} &\approx \xi_{10} + \varepsilon_1 \xi_{00} \\ &\approx \tau_1 \dot{s}(t) + \varepsilon_1 s(t).\end{aligned}\quad (12)$$

From second-order statistics, the correlation between any signal and its time derivative is zero

$$E[s(t)\dot{s}(t)] = 0 \quad (13)$$

resulting in zero correlation between common-mode and gradient components

$$E[\xi_{00}\xi_{10}] = 0. \quad (14)$$

Applying (14) to (12) yields an improved estimate of the spatial gradient that eliminates the systematic common-mode error

$$\xi_{10} \approx \hat{\xi}_{10} - \frac{E[\xi_{00}\hat{\xi}_{10}]}{E[\xi_{00}^2]} \xi_{00}. \quad (15)$$

To efficiently implement the correction factor  $\varepsilon_1$  in the digital domain, a sign-sign LMS (SS-LMS) adaptation rule is used [23].  $\varepsilon_1$  is stored as digital value in a 12-bit counter and it is represented in two's complement. The update is performed by incrementing or decrementing the counter based on the polarities of spatial gradient and common-mode signals

$$\begin{aligned}\varepsilon_1^+[n+1] &= \varepsilon_1^+[n] + \text{sgn}(\xi_{10}^+[n] - \xi_{10}^-[n]) \\ &\quad \times \text{sgn}(\xi_{00}^+[n] - \xi_{00}^-[n]) \\ \varepsilon_1^-[n+1] &= 2^{12} - 1 - \varepsilon_1^+[n+1].\end{aligned}\quad (16)$$

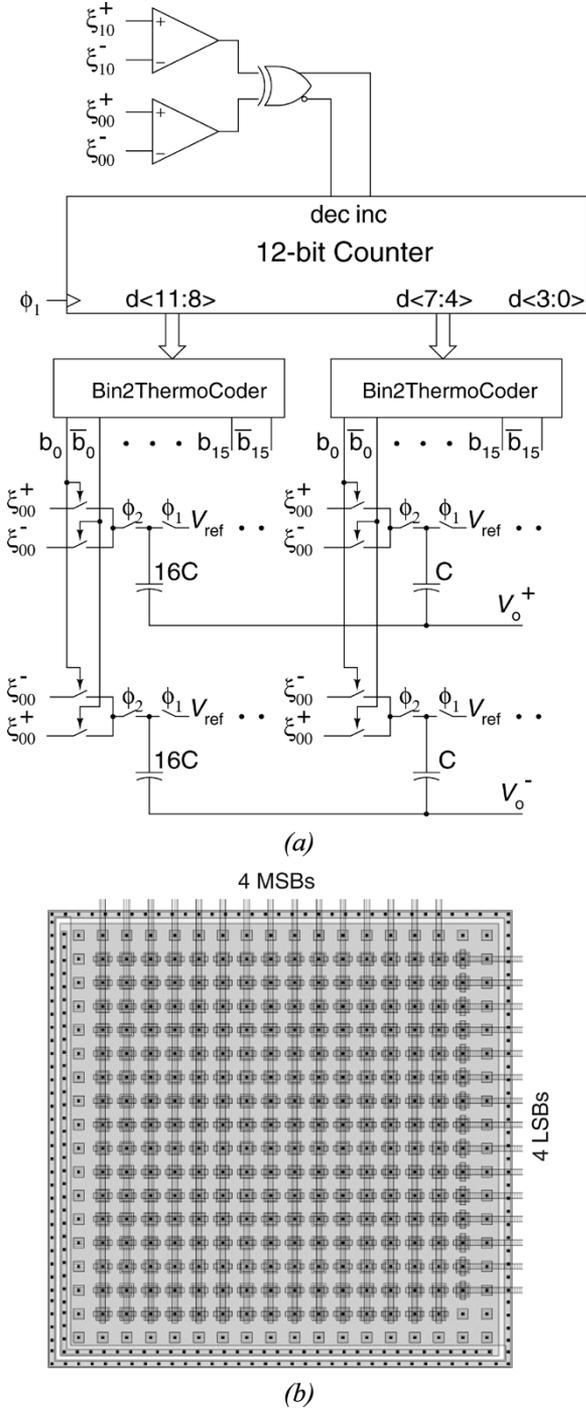


Fig. 6. (a) Implementation of SS-LMS adaptation rule and multiplying D/A capacitor arrays connecting to nodes  $V_o^+$  and  $V_o^-$  in Fig. 5. (b) Layout of  $16 \times 16$  capacitor array.

The 8 most significant bits (MSBs) are presented to multiplying digital-analog (D/A) capacitor arrays to construct the LMS error signal and the correction term  $\varepsilon_1 \xi_{00}$ . Implementation of the SS-LMS adaptation rule and capacitor arrays is shown in Fig. 6(a). Capacitive coupling of the D/A arrays to nodes  $V_o^+$  and  $V_o^-$  in Fig. 5(b) establishes the corrected gradient estimate

$$\begin{aligned} \xi_{10}^+[n] &= \hat{\xi}_{10}^+[n] - (\varepsilon_1^+ \xi_{00}^+[n] + \varepsilon_1^- \xi_{00}^-[n]) \\ \xi_{10}^-[n] &= \hat{\xi}_{10}^-[n] - (\varepsilon_1^- \xi_{00}^+[n] + \varepsilon_1^+ \xi_{00}^-[n]). \end{aligned} \quad (17)$$

In phase  $\phi_1$ , both capacitor arrays are precharged to zero-level reference voltage  $V_{ref}$ . In phase  $\phi_2$ , a fraction of the capacitance corresponding to  $\varepsilon_1^+$  (a number of unit capacitors equal to the 8-bit representation) couple  $\xi_{00}^+$  to the branch  $\xi_{10}^+$ , while the complement (the remainder of the capacitor array) couple  $\xi_{00}^-$  to  $\xi_{10}^+$ . To implement the complement  $\varepsilon_1^-$  in branch  $\xi_{10}^-$ , the connections to  $\xi_{00}^+$  and  $\xi_{00}^-$  in the capacitor array are exactly reversed. The additional 4 bits of stored  $\varepsilon_1$  value provide a digital buffer integrating LMS contributions that offers flexibility in the choice of learning rate, regulated by the frequency of updating.

The capacitor array is implemented as an array of 15 columns each containing 16 unit capacitors, controlled by the 4 MSBs presented in thermometer code, and a column of 15 identical unit capacitors controlled by the next 4 least significant bits (LSBs) in thermometer code. Layout of the array is depicted in Fig. 6(b). To reduce the size of the array, bottom plate sampling was not implemented. The bottom plate common for all unit capacitors is connected to the amplifier's virtual ground node. The value of unit capacitor is 5.1 fF.

The range of common-mode error correction is designed to compensate for at most 5 percent of common-mode leakage in the spatial gradient estimates. Since the minimum size of capacitor array is constrained by the size of unit capacitor, a T-cell is used to attenuate the output swing of the multiplying D/A capacitor array. While the T-cell leads to nonlinearity in the transfer characteristic, it does not affect the monotonicity of the D/A which is sufficient for LMS convergence. Stray sensitivity in the T-cell does not influence the differential linearity. The values of capacitors  $C_3$  and  $C_4$  are 1 and 0.1 pF, respectively.

### C. Bearing Estimation

Estimation of bearing ITDs is implemented with digital SS-LMS differential on-line adaptation, using architecture similar to common-mode error correction. Bearing estimates are represented as 12-bit values in two's complement

$$\begin{aligned} \tau_1^+[n+1] &= \tau_1^+[n] + \text{sgn}(e_{10}^+[n] - e_{10}^-[n]) \\ &\quad \times \text{sgn}(\xi_{10}^+[n] - \xi_{10}^-[n]) \\ \tau_1^-[n+1] &= 2^{12} - 1 - \tau_1^+[n+1] \end{aligned} \quad (18)$$

where the 8 MSBs are used for computation of LMS error signal

$$\begin{aligned} e_{10}^+[n] &= \xi_{10}^+[n] - (\tau_1^+ \dot{\xi}_{00}^+[n] + \tau_1^- \dot{\xi}_{00}^-[n]) \\ e_{10}^-[n] &= \xi_{10}^-[n] - (\tau_1^- \dot{\xi}_{00}^+[n] + \tau_1^+ \dot{\xi}_{00}^-[n]). \end{aligned} \quad (19)$$

The SC realization is shown in Fig. 7. Clock  $\phi_{2e}$  represents a replica of  $\phi_2$  with falling edge slightly preceding the falling edge of  $\phi_2$ . The SS-LMS adaptation stage and multiplying D/A capacitor array are identical to the ones described in Section III-B (Fig. 6), but with clock phases  $\phi_1$  and  $\phi_2$  reversed accounting for a half clock cycle delay between gradient computation and bearing estimation stages.

With the given capacitor sizing, the minimum observable ITD delay relative to the sampling interval  $T$  equals  $T/640$ , and the maximum observable ITD delay equals  $2/5T$ . The number of update iterations (18) to convergence depends on the signal and by virtue of the SS-LMS rule does not exceed  $2^{12}$  update cycles, determined by the register length.

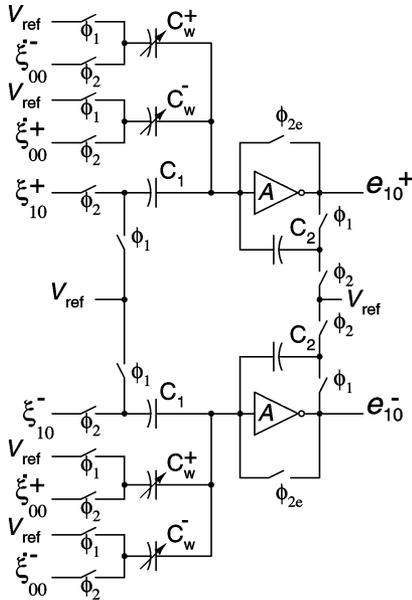


Fig. 7. SC realization of SS-LMS adaptive ITD delay estimation.

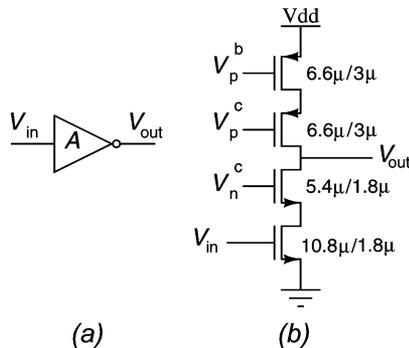


Fig. 8. Cascoded nMOS inverting amplifier. (a) Symbol. (b) Circuit diagram.

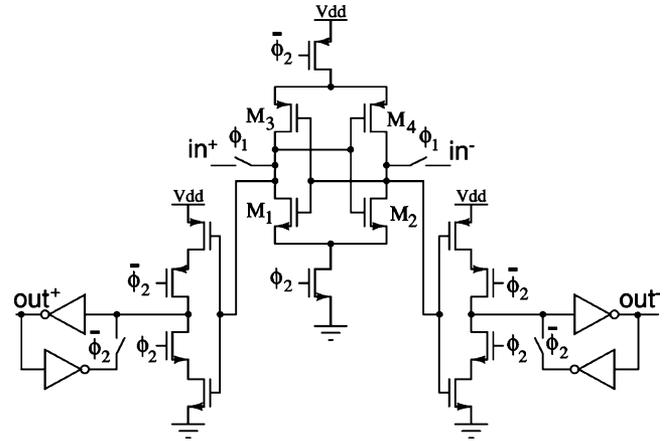
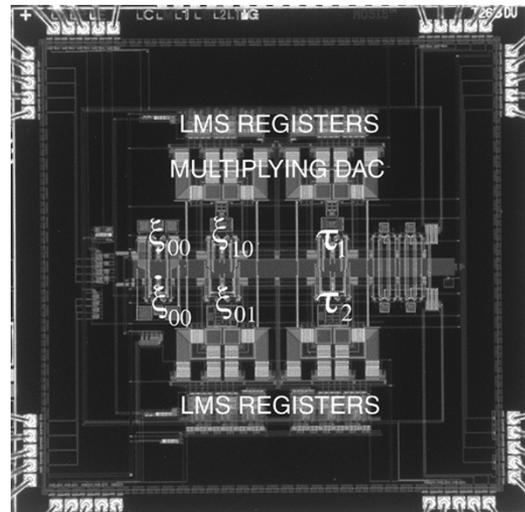
#### D. Amplifier Design

A cascoded pseudo-nMOS inverter, shown in Fig. 8, is used as gain element in the above fully differential SC circuits. The choice of telescopic operational amplifier without tail transistor results in smaller area and reduced noise and power dissipation [24], [25]. The most energy-efficient solution for low-frequency applications is to bias the amplifier in weak inversion (the upper range of the subthreshold regime), where the amplifier has maximum transconductance-to-current ratio for maximum energy efficiency at the highest possible speed, while providing extended output dynamic range at low voltage supply [26].

The voltage bias  $V_p^b$  controls the biasing current of the amplifier, set by considerations of sampling frequency, slew-rate, and power dissipation. At 200 nA of biasing current, 1-pF load capacitance and 3-V supply, simulations indicate an open-loop dc gain of 91 dB and gain-bandwidth product of 844.3 kHz. No additional gain-boosting techniques were attempted, since the provided dc gain was sufficient for the desired resolution.

#### E. Comparator Design

The sign function in the SS-LMS update for bearing estimation (18) is implemented by a latched, regenerative com-

Fig. 9. Comparator circuit diagram. Transistors M1 through M4 are sized  $3 \mu/1.2 \mu$ , and all other transistors are minimum size.Fig. 10. Micrograph of  $3 \text{ mm} \times 3 \text{ mm}$  chip in  $0.5\text{-}\mu\text{m}$  CMOS technology.

parator shown in Fig. 9. For the common-mode suppression SS-LMS updates (16), the comparator is implemented by reversing the clock phases  $\phi_1$  and  $\phi_2$  in Fig. 9, accounting for the half clock cycle delay between the common-mode suppression and bearing estimation SS-LMS stages (see Section III-C).

The offset of the comparator was measured to be on the order of 10 mV. Kick-back noise inherent to this comparator implementation forces alternating comparison results in consecutive cycles when the input difference is near zero, countering the effect of offset on SS-LMS adaptation at convergence.

#### IV. EXPERIMENTAL RESULTS

The VLSI acoustic localizer measures  $3 \text{ mm} \times 3 \text{ mm}$  in  $0.5\text{-}\mu\text{m}$  CMOS technology. Fig. 10 shows the micrograph of the chip, highlighting the functional blocks as described in Section III.

All experiments were conducted at 3-V supply voltage, with the zero-level reference  $V_{\text{ref}}$  set to 1.5 V. The signal swing of the sensor input is 600 mV peak-to-peak. The single-sided signal swings at the output of inverters are optimized to be 2.4 V peak-to-peak, except for the common-mode signals  $\xi_{00}^+$

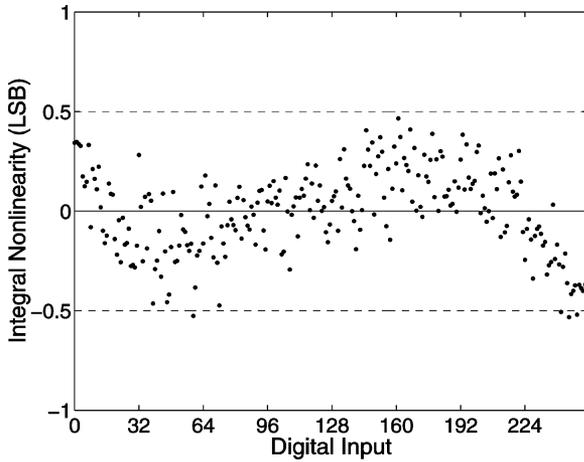


Fig. 11. INL of multiplying D/A converter.

and  $\xi_{00}^-$ , which are used in common-mode suppression and further attenuated by the T-cells.

The digital ITD estimates  $\tau_1$  and  $\tau_2$ , obtained directly from the bearing registers (18) at convergence, are output in bit-serial format using a separate clock. The chip also outputs the gradient signals  $\xi_{10}$ ,  $\xi_{01}$  and  $\xi_{00}$ , for use in separation and localization of multiple (up to three) acoustic sources [20]. The gradient output signals are presented in complementary analog format through sample-and-hold buffers.

#### A. Chip Characterization

For characterization, the chip was supplied with synthetic input signals bypassing the microphones. In the first experiment, all input signals  $x_{10}$ ,  $x_{01}$ ,  $x_{-10}$  and  $x_{0-1}$  were connected together and a two-level alternating signal was presented. This emulated the condition of zero spatial gradients, and constant temporal derivative. The content of the ITD register in the SS-LMS bearing estimation stage was varied from 0 to 255, and the differential error signal of the corresponding bearing estimator ( $e_{10}^+ - e_{10}^-$ ) was recorded as displayed in Fig. 11. The resulting measure of integral nonlinearity (INL) of the multiplying D/A capacitor array includes errors contributed by the gradient computation and bearing estimation stages, which account for the observed curvature in Fig. 11.

System-level accuracy of ITD delay estimation was demonstrated using a 200 Hz synthetic sine wave input signal, characteristic of acoustic signatures of ground vehicles, and a system sampling frequency of 2 kHz. The signal was presented to both  $x_{10}$  and  $x_{01}$ , and with variable delay to both  $x_{-10}$  and  $x_{0-1}$ . Fig. 12 illustrates computation of the spatial gradient for a given delay. Digital bearing estimates were recorded with common-mode suppression activated. The delay was varied from  $-400 \mu\text{s}$  to  $400 \mu\text{s}$ , in  $2 \mu\text{s}$  increments. A 2 s lag in the recording ensured full convergence of common-mode and bearing SS-LMS adaptation ( $2^{12}$  update cycles). The recorded digital ITD delay estimate  $\tau_1$  as a function of actual delay is shown in Fig. 13.

Note the super-resolution nature of the delay estimation, with  $2\text{-}\mu\text{s}$  time resolution at 2-kHz sampling rate. The ability to resolve delays that are a small fraction (0.4%) of the sampling interval is characteristic of gradient flow, and allows to reduce

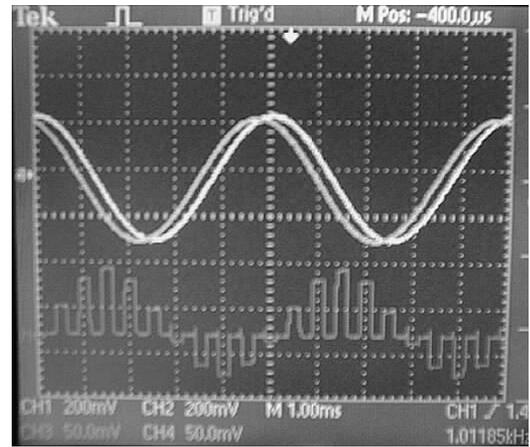


Fig. 12. Estimation of spatial gradient. *Top*: delayed sine wave inputs  $x_{10}$  and  $x_{-10}$ . *Bottom*: differenced gradient signal  $\xi_{10} = \xi_{10}^+ - \xi_{10}^-$ .

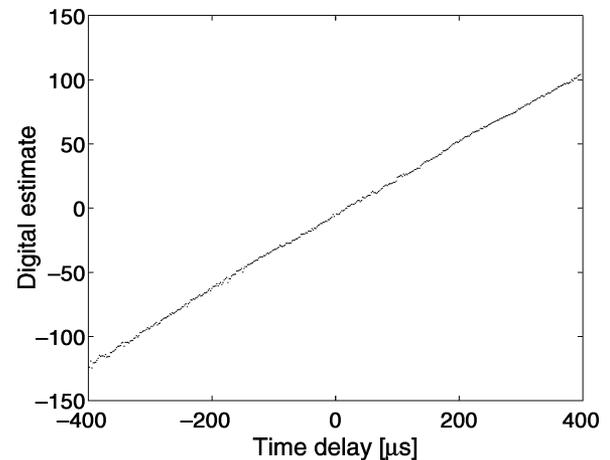


Fig. 13. Digital estimation of ITD delay  $\tau_1$  for synthetic 200 Hz sine wave signals, sampled at 2 kHz.

the power. At the 2-kHz clock (suitable for acoustic surveillance applications), power consumption measured  $30 \mu\text{W}$  from the analog (3 V) supply, and  $1.8 \mu\text{W}$  from the digital supply, for a total of  $32 \mu\text{W}$ . At 16 kHz clock (for hearing aid and audio-range multimedia applications) measured power increased to  $54 \mu\text{W}$ . About 30% of these figures account for the power consumed by the sample-and-hold output buffers, not used in localization but to observe the gradient signals.

#### B. Localization Performance

To evaluate localization performance in the audio range, the chip was interfaced with a 1 cm radius microphone array of the topology shown in Fig. 1(a). A single acoustic source was presented through a loudspeaker positioned at 1-m distance from the array. For precise control of angle  $\theta$  in the experiments, the platform containing the array and chip was rotated, using a stepper motor, around the axis through the center of the array perpendicular to the plane.

The planar array of four omnidirectional hearing aid miniature microphones (Knowles IM-3268) is shown in the inset of Fig. 14(a). Opposing pairs of microphones in the array are oriented identically, rather than symmetrically, to minimize differences in microphone response due to mismatch in directionality

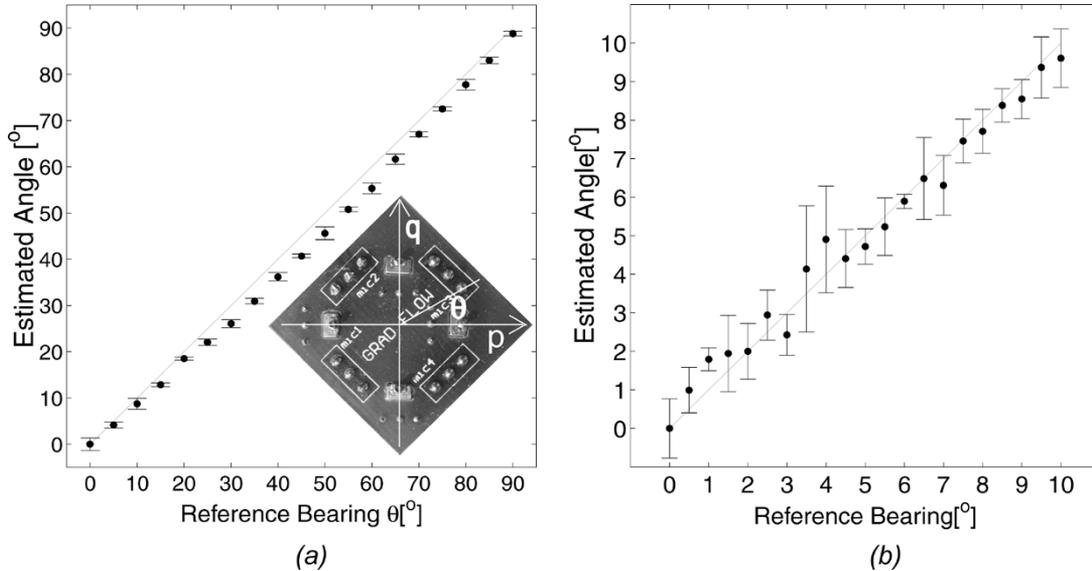


Fig. 14. Localization of bandlimited (100–1000 Hz) Gaussian source. (a) 5 degree increments. (b) 0.5 degree increments.

pattern. The common-mode suppression (Section III-B) compensates for up to 5% in gain mismatch between microphones. Phase mismatch is unaccounted for, except for linear phase errors (pure delay differences) which introduce offset in ITD estimation that can be digitally subtracted (not implemented).

The system sampling frequency of the chip was set to 16 kHz. A broad-band bandlimited (100–1000 Hz) Gaussian signal was presented. The volume was adjusted for the microphones to supply full-scale inputs to the chip (200 mV). The corresponding signal-to-noise (SNR) ratio was around 35 dB. Expressions for Cramer–Rao lower bound on variance of bearing angle estimate [1] predict an achievable angular resolution of around 1 degree under these conditions and a 1 s observation interval.

For each setting of bearing angle, 10 samples of ITD estimates were recorded from the chip at 1 s intervals. The platform was rotated in increments of  $5^\circ$  until it reached  $90^\circ$  from the initial position. The absolute angle  $\theta$  between speaker and microphone array ranged from around  $70^\circ$  to  $-20^\circ$ . The mean and variance of estimated bearing angles normalized from the initial position are shown in Fig. 14(a). To assess finer scale resolution, the bearing angle was then varied from  $0^\circ$  to  $10^\circ$  in increments of  $0.5^\circ$  corresponding to a range of absolute angles  $\theta$  between speaker and array from  $50^\circ$  to  $40^\circ$ . Fig. 14(b) presents the mean and variance of the bearing angle estimates from recorded ITDs.

### C. Open-Field Vehicle Tracking

After assessing the localization performance in controlled room environment experiments, tests were conducted in the open-field with the chip tracking different types of ground-based vehicles, plus some occasional airborne traffic. For these experiments, the acoustic localizer chip was integrated into an acoustic surveillance unit (ASU) enclosure, designed and implemented by Signal Systems Corporation (SSC) [27]. Measuring 11 cm in diameter, the ASU also contained four Knowles SiSonic MEMS microphones and signal conditioning circuitry.

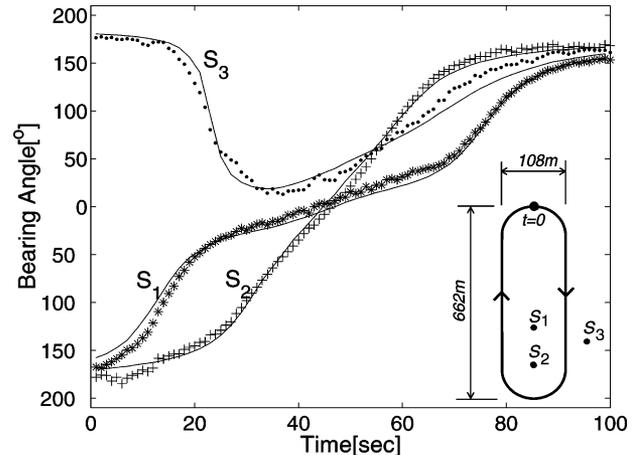


Fig. 15. Experimental tracking of a ground vehicle using three acoustic localization sensor nodes positioned in the field as shown in the inset. Solid lines indicate ground truth estimates from GPS measurement.

While the microphones were closely spaced in the center, provisions for wind-noise reduction increased the effective aperture to the diameter of the enclosure. The larger 11 cm aperture is desirable for the frequency range of interest with vehicles emitting motor noise in the 100–300 Hz range. To accommodate this range the sampling clock frequency of the chip was set to 2 kHz. The ASU, including the gradient flow localization chip, was powered with a chain of 3 AA batteries without voltage regulator.

Ground-based vehicles were driven around a  $662\text{ m} \times 108\text{ m}$  oval-shaped track at different speeds, and tracked with three ASUs positioned at points  $S_1$ ,  $S_2$  and  $S_3$  approximately 90 m apart as illustrated in Fig. 15. The tests were performed with relatively loud ambient background noise. The range at which the localizer was able to track the vehicles depended on the type of vehicle and varied from 350 m to 600 m, limited by test conditions. As an example, tracking of one vehicle moving clockwise around the oval is illustrated in Fig. 15. The bearing ITD estimates were recorded at 1 s time intervals and converted to

TABLE I  
ACOUSTIC LOCALIZER CHARACTERISTICS

Technology	0.5 $\mu\text{m}$ 2P3M CMOS
Size	3 mm $\times$ 3 mm
Supply	3 V
Resolution	2 $\mu\text{s}$ at 2 kHz 0.25 $\mu\text{s}$ at 16 kHz
Power dissipation	32 $\mu\text{W}$ at 2 kHz 54 $\mu\text{W}$ at 16 kHz

angular form using (3). Only the azimuth angles  $\theta$  are shown for ground vehicles, since the ASU microphone arrays were oriented horizontally. Estimates of true bearing angles  $\theta$  from GPS tracking of the vehicles, accounting for the approximate geometry of the track and ASU's and correcting for delays in acoustic wave propagation, are also depicted as "ground truth" in Fig. 15. In tests with multiple vehicles moving together or in opposite directions, the localizer chip tracked the loudest (received) target. A F/A-18 fighter jet flying overhead at approximately 3000 m was also successfully tracked by the localizer chip, accounting for elevation besides azimuth.

## V. CONCLUSION

A micropower adaptive VLSI chip for gradient-based bearing estimation in miniature acoustic arrays was presented, that combats problems of gain mismatch and  $1/f$  noise originating from the sensor array and preamplifiers. The obtained time-resolution leads to fine angular resolutions in localization of acoustic sources. The ability to track ground-based and airborne vehicles was experimentally verified in field tests. The measured characteristics are summarized in Table I. The chip is suitable for integration with a MEMS microphone array, for applications in surveillance and hearing aids.

Besides acoustics, the gradient flow principle extends to EM and other traveling wave sensing modalities. Interestingly, wavelengths of EM radiation in the RF frequency range coincide with those for sound propagation in the audio range, and both are amenable for sensor integration on a single silicon die. We anticipate that application of gradient flow to RF localization will enable single-chip sub-ps bearing resolution at sampling rates determined by signal bandwidth, and not the RF carrier.

## ACKNOWLEDGMENT

The authors would like to thank Abdullah Celik, Larry Riddle and Signal Systems Corporation for measurement help.

## REFERENCES

- [1] B. Friedlander, "On the Cramer-Rao bound for time delay and Doppler estimation," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 3, pp. 575–580, Mar. 1984.
- [2] J. Lazzaro and C. A. Mead, "A silicon model of auditory localization," *Neur. Comput.*, vol. 1, pp. 47–57, 1989.
- [3] N. Bhadkamkar and B. Fowler, "A sound localization system based on biological analogy," in *Prpc. IEEE Int. Conf. on Neural Networks*, vol. 3, 1993, pp. 1902–1907.
- [4] T. Horiuchi, "An auditory localization and coordinate transform chip," in *Advances in Neural Information Processing Systems*. New York: Morgan Kaufmann, 1995, vol. 7, pp. 787–794.
- [5] J. G. Harris, C.-J. Pu, and J. C. Principe, "A Neuromorphic Monaural Sound Localizer," in *Advances in Neural Information Processing Systems*. New York: Morgan Kaufmann, 1999, vol. 11.
- [6] I. Grech, J. Micallef, and T. Vladimirova, "Experimental results obtained from analog chips used for extracting sound localization cues," in *Proc. 9th Int. Conf. Electronics, Circuits and Systems*, vol. 1, Sept. 15–18, 2002, pp. 247–251.
- [7] A. van Schaik and S. Shamma, "A neuromorphic sound localizer for a smart MEMS system," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'2003)*, Bangkok, Thailand, May 25–28, 2003.
- [8] P. Julian, A. G. Andreou, P. Mandolesi, and D. Goldberg, "A low-power CMOS integrated circuit for bearing estimation," in *Proc. IEEE Int. Symp. on Circuits and Systems (ISCAS'2003)*, Bangkok, Thailand, May 25–28, 2003.
- [9] J. C. Chen, L. Yip, J. Elson, H. Wang, D. Maniezzo, R. Hudson, K. Yao, and D. Estrin, "Coherent acoustic array processing and localization on wireless sensor networks," *Proc. IEEE*, vol. 91, no. 8, pp. 1154–1162, Aug. 2003.
- [10] R. J. Kozick and B. M. Sadler, "Source localization with distributed sensor arrays and partial spatial coherence," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 601–616, 2004.
- [11] P. Aarabi, "The fusion of distributed microphone arrays for sound localization," *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 4, pp. 338–347, 2003.
- [12] N. Strobel, S. Spors, and R. Rabenstein, "Joint audio-video object localization and tracking," *IEEE Signal Proc. Mag.*, vol. 18, no. 1, pp. 22–31, Jan. 2001.
- [13] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [14] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-24, no. 11, pp. 320–327, Nov. 1976.
- [15] P. Julian, A. G. Andreou, L. Riddle, S. Shamma, D. H. Goldberg, and G. Cauwenberghs, "Comparative study of sound localization algorithms for energy aware sensor network nodes," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 51, no. 4, pp. 640–648, Apr. 2004.
- [16] J. Barrère and G. Chabriel, "A compact sensor array for blind separation of sources," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 5, pp. 565–574, May 2002.
- [17] G. Cauwenberghs, M. Stanacevic, and G. Zweig, "Blind broad-band source localization and separation in miniature sensor arrays," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'2001)*, Sydney, Australia, May 6–9, 2001.
- [18] D. Robert, R. N. Miles, and R. R. Hoy, "Tympanal hearing in the sarcophagid parasitoid fly *emblemiasoma* sp.: the biomechanics of directional hearing," *J. Exper. Biol.*, vol. 202, pp. 1865–1876, 1999.
- [19] R. Reeve and B. Webb, "New neural circuits for robot phonotaxis," *Phil. Trans. Roy. Soc. A*, vol. 361, pp. 2245–2266, 2002.
- [20] A. Celik, M. Stanacevic, and G. Cauwenberghs, "Mixed-signal real-time adaptive blind source separation," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'2004)*, Vancouver, BC, Canada, May 23–26, 2004.
- [21] M. Stanacevic and G. Cauwenberghs, "Micropower mixed-signal acoustic localizer," in *Proc. IEEE Eur. Solid State Circuits Conf. (ESSCIRC'2003)*, Estoril, Portugal, Sep. 16–18, 2003.
- [22] C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization," *Proc. IEEE*, vol. 84, no. 11, pp. 1584–1614, Nov. 1996.
- [23] L. Der and B. Razavi, "A 2-GHz CMOS image-reject receiver with LMS calibration," *IEEE J. Solid-State Circuits*, vol. 38, no. 2, pp. 167–175, Feb. 2003.
- [24] G. Nicollini, F. Moretti, and M. Conti, "High-frequency fully differential filter usign operational amplifier without common-mode feedback," *IEEE J. Solid-State Circuits*, vol. 24, no. 3, pp. 803–813, Mar. 1989.
- [25] J. Li and U. Moon, "A 1.8-V 67-mW 10-bit 100-MS/s pipelined ADC using time-shifted CDS technique," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1468–1476, Sep. 2004.
- [26] E. Vittoz, "Micropower techniques," in *Design of Analog-Digital VLSI Circuits for Telecommunications and Signal Processing*, 2nd ed, J. Franca and Y. Tsvividis, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1994, pp. 53–96.
- [27] L. Riddle *et al.*, "VLSI acoustic surveillance unit," in *Proc. GOMAC*, Monterey, CA, Mar. 15–18, 2004.



**Milutin Stanaćević** (SM'00) received the B.S. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia and Montenegro, in 1999, and the M.S. degree in electrical and computer engineering from Johns Hopkins University, Baltimore, MD, in 2001. He is currently working toward the Ph.D. degree at Johns Hopkins University.

His research interests include mixed-signal VLSI circuits, systems, and algorithms for parallel multi-channel sensory information processing with emphasis on real-time acoustic source localization and separation, and micropower implantable biomedical instrumentation and telemetry.



**Gert Cauwenberghs** (SM'89–M'94–S'03) received the Ph.D. degree in electrical engineering from the California Institute of Technology in 1994, and is presently Professor of Electrical and Computer Engineering at Johns Hopkins University, Baltimore.

His research covers VLSI circuits, systems and algorithms for parallel signal processing, adaptive neural computation, adaptive optics, and biomedical instrumentation.

He received the National Science Foundation Career Award in 1997, the Office of Naval Research Young Investigator Award in 1999, and the White House PECASE Award in 2000. He serves as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I : REGULAR PAPERS, and the IEEE SENSORS JOURNAL. He was a Distinguished Lecturer of the IEEE Circuits and Systems Society and chaired its Analog Signal Processing Technical Committee.