# The system-on-chip design of a silicon micromachined condenser microphone array

Mingsian R. Bai[a] and Shihwei Huang
*Department of Mechanical Engineering, National Chiao-Tung University, 1001 Ta-Hsueh Road, Hsin-Chu 300, Taiwan, Republic of China*

(Received 29 August 2003; revised 25 March 2004; accepted 30 March 2004)

A miniature condenser microphone array that combines array signal processing and micro-electro-mechanical systems (MEMS) technologies is presented. A linear dynamic model and a quasistatic analysis are presented. The array configuration serves two purposes: enhancement of the signal-to-noise ratio (SNR) and the directivity of sensor. A least-squares beamforming design based on the template resulting from the Multiple Signals Classification algorithm is proposed to achieve a directive beam pattern. To minimize the parasitic effects on the MEMS device, a system-on-chip design composed of a microphone module, a dc bias circuit, an impedance matching circuit, and array beamforming filters is proposed in the paper. The performance of the proposed design is evaluated in terms of the frequency response, SNR improvement, and directional response through numerical simulations and experiments. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1753295]

PACS numbers: 43.38.Bs, 43.38.Gy, 43.38.Hz, 43.38.Kb [AJZ] Pages: 303–312

## I. INTRODUCTION

3C (computer, communication, and consumer) industry has recently emerged as one of the most rapidly growing industries of modern days. Commercial products such as notebook computers, personal data assistants, mobile phones, and MP3 players are currently making their way into people's lives. Miniaturization has been known to be one of the chief concerns of 3C products. This is particularly true for microphones that are important components of 3C products. A condenser microphone based on micro-electro-mechanical systems (MEMS) technology offers a potential solution in achieving ultimate miniaturization for 3C products. Silicon microphones have received a great deal of research interest for years and an excellent review can be found in Ref. 1. However, there remain technical issues that need to be resolved before we find widespread use of such device. Although conventional capacitive microphones have higher sensitivity than piezoresistive microphones, sensitivity is still a crucial consideration in micromachined capacitive microphones. This sensitivity problem pertinent to silicon microphones is due partly to the built-in stress in the diaphragm resulting from the MEMS fabrication process, and partly to the parasitic capacitance in electrical connections. This problem is further worsened by the self-noise in the chip, which may overwhelm the exceedingly weak signal generated by the device. Hence, the signal-to-noise ratio (SNR) has always been one of the major design factors in MEMS microphones.

To address this issue of low sensitivity, a system-on-chip (SoC) design of the microphone that integrates the sensor structure, a dc bias circuit, an impedance buffer, and array beamforming filters in one chip is presented in this paper. The purpose of this paper is threefold. The first purpose is to show how electroacoustic transducers, array signal processing, and MEMS technology can be applied in a multidisciplinary design problem. The second purpose is to improve SNR through integration into arrays. The third purpose is to improve directionality through integration into arrays. However, for the last purpose, it shall be shown that it can be achieved above moderately high frequency due to the physical constraint of array size. In the single chip design, parasitic effects are reduced due to elimination of interconnections between modules. Apart from the SoC design, a microphone array configuration is also proposed in this study to further enhance the SNR of the device. One benefit of using such array configuration is that the SNR (in terms of power) increases by the number of elements in the array.[2] Furthermore, the microphone array serves as a spatial filter to focus only on the source and reject unwanted noise at the other directions, making the array an attractive solution for hands-free communications. In order to achieve a highly directive beam pattern, a constant beamwidth array based on the template from the Multiple Signals Classification (MUSIC)[3] algorithm is designed using the least-squares method. For simplicity, the resulting filters are realized by analog circuits. The MEMS fabrication is well suited for such multichannel array SoC design.

In this paper, the properties of the condenser microphone and array filters are investigated. Electro-acoustical analogy is employed to establish a linear dynamic model of the silicon microphone, while a quasistatic analysis based on the finite difference method is conducted to find the collapse condition. Some previous research on this aspect should be mentioned. Hohm and Kühnel[4,5] applied an analytical approach to calculate the nonlinear deflection of the diaphragm of a silicon microphone. Bergqvist[6] suggested a more complex model based on the finite elements method (FEM). Pedersen[1] examined the effects of the perforation in the back plate on microphone stability and performance by using the

[a]Author to whom all correspondence should be addressed; electronic mail: msbai@mail.nctu.edu.tw

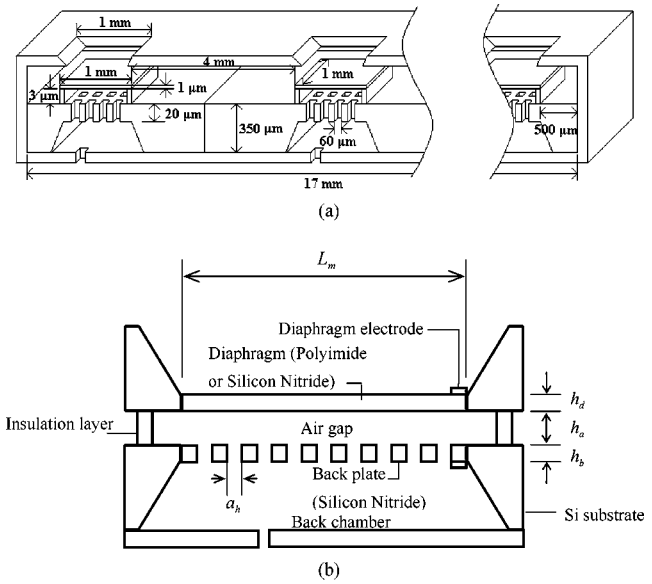


FIG. 1. A schematic diagram of the microphone array. (a) The microphone array with four MEMS condenser microphones. (b) Cross section of a condenser microphone with a perforated backplate.

finite differences method (FDM). Similar to the work by Pedersen and other researchers,[4–6] electro-acoustical analogy[7–9] is exploited in this paper to account for the complex coupling among the acoustical, mechanical, and electrical domains. The collapse condition of the microphone is determined from a quasistatic analysis of the nonlinear system. On the basis of the linear dynamic model, performance analysis was carried out to justify the proposed MEMS microphone array. As for the MEMS microphone arrays, two references are relevant to the present study. Chowdhury *et al.* presented a 3×3 MEMS capacitive microphone array.[10] In their design, beamforming was accomplished by the simple delay-sum method. However, there was an error in the analysis of directional response, and no experimental results are presented. Another work by Arnold *et al.* developed a 16-element MEMS piezoresistive microphone array for wind tunnel studies.[11] The fast Fourier transform is used for frequency-domain beamforming. However, instead of being fabricated on one chip, the microphones are mounted on a printed circuit board. Thus, the performance of such design was found comparable to conventional microphone arrays.

## II. DESIGN OF MEMS CONDENSER MICROPHONE ARRAY

Consider a four-channel MEMS microphone array shown in Fig. 1(a). Four condenser microphones are fabricated with equal spacing on a straight line. In this section, the quasistatic and the linear dynamic models and the array beamforming design will be presented.

### A. Quasistatic analysis and linear dynamic analysis

We begin with a single channel condenser microphone with a square diaphragm and a perforated backplate, as shown in Fig. 1(b). In the metallization process, Cr and Au are deposited using evaporation as the adhesion layer and the electrode layer, respectively. The metallization layer can be released by a lift-off process. The metallization scheme should play a role in the diaphragm stress and the parasitic capacitance. Like traditional condenser microphones, a dc bias voltage source is needed in our MEMS microphone. To predict the large deflection of the diaphragm resulting from an excessive dc bias, a quasistatic analysis in Ref. 1 is reviewed.

Assume that the backplate is rigid and the deflection is small so that the linear model applies. For simplicity, nonlinear stiffening that may affect the prediction of the electrostatic collapse point is not considered in the following presentation. The equation of motion for the diaphragm is given as[12]

$$C_{11}\frac{\partial^4 w_d}{\partial x^4}+2(C_{12}+2C_{44})\frac{\partial^4 w_d}{\partial x^2\partial y^2}+C_{11}\frac{\partial^4 w_d}{\partial y^4}=\frac{12}{h_d^3}\left[p_{\mathrm{sp}}+p_{\mathrm{el}}+\sigma_d h_d\left(\frac{\partial^2 w_d}{\partial x^2}+\frac{\partial^2 w_d}{\partial y^2}\right)\right], \tag{1}$$

where $h_d$ is the plate thickness, $\sigma_d$ is the built-in stress, and $w_d(x,y)$ is the deflection. $C_{11}=C_{22}$, $C_{12}=C_{21}$, and $C_{44}$ are material constants of the plate. For isotropic materials, the forgoing material constants are given by

$$C_{11}=\frac{E}{1-\upsilon^2},\quad C_{12}=\frac{E\upsilon}{1-\upsilon^2},\quad C_{44}=\frac{E}{2(1+\upsilon)}, \tag{2}$$

where $E$ and $\upsilon$ are the Young's modulus and Poisson's ratio, respectively. In Eq. (1), $p_{\mathrm{sp}}$ is the sound pressure, $p_{\mathrm{el}}$ represents the electrostatic force per unit area due to the dc bias between the diaphragm and the backplate and is given by

$$p_{\mathrm{el}}(x,y)=K_{\mathrm{holes}}\frac{\epsilon_d\epsilon_0}{2(h_d+\epsilon_d(h_a-w_d))}V_{ba}^2, \tag{3}$$

$\epsilon_d$ is the relative permittivity of the diaphragm material, $\epsilon_0$ is the vacuum permittivity, $h_a$ is the distance of air gap, and $V_{ba}$ is the dc bias voltage. The constant $K_{\mathrm{holes}}$ accounts for the effect when the backplate is perforated.[1] Although in a practical setting the diaphragm is fixed by a tapered and somewhat asymmetric junction, it is assumed for convenience that the diaphragm is clamped at the edge, i.e.,

$$w_d(x,y)=\frac{\partial w_d(x,y)}{\partial x}=0,\quad w_d(x,y)=\frac{\partial w_d(x,y)}{\partial y}=0$$

on the boundary. An iterative quasistatic analysis based on the FDM is carried out for the partial differential equations in Eq. (1). The large deflection of diaphragm and the collapse condition of the microphone is found from this analysis. The details of the procedure can be found in Ref. 1.

Next, electro-acoustical analogy[1,8,9] is adopted for predicting the linear dynamic behavior of the MEMS condenser microphone. The equivalent circuits of the microphone are shown in Fig. 2, wherein the acoustical, mechanical, and electrical domains are coupled through ideal transformers.

In the acoustical subsystem of Fig. 2(b), the radiation impedance is approximated by the analogous circuits[8] in Fig. 2(c), where $R_{A1}=0.441\rho_0 c/L_m^2$, $C_{A1}=5.94L_m^3/\rho_0 c^2\pi\sqrt{\pi}$, $R_{A2}=\rho_0 c/L_m^2$, $M_{A1}=8\rho_0/3\pi\sqrt{\pi}L_m$, $\rho_0$ is the density of air,

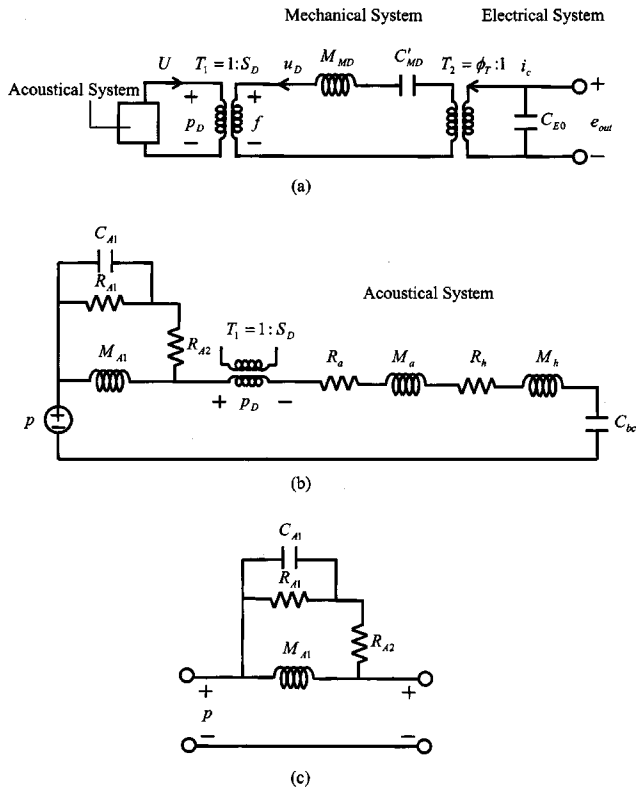


FIG. 2. The equivalent circuit based on electro-acoustical analogy of the condenser microphone. (a) Complete system composed of three coupled subsystems: the acoustical system, mechanical system, and electrical system. (b) Detailed circuit representation of the acoustical subsystem. (c) Circuit representation of the radiation impedances.

$c$ is the sound speed in air, and $L_m$ is the width of the square diaphragm. Similar to the approach taken by Skvor,[13] the effect of the air gap is modeled as equivalent resistance and mass elements in the acoustical domain: $R_a = (1.22\eta\pi b^2/h_a^3 L_m^2)B$ and $M_a = (0.102\rho_0\pi b^2/h_a L_m^2)B$, where $\eta = 1.86\times 10^{-5}$ N s/m$^2$ at 20 °C is the dynamic viscosity of air, $h_a$ is the distance of the air gap, and $B$ is defined as

$$B = \frac{1}{4}\ln\left(\frac{0.16b^2}{a_h^2}\right) - \frac{3}{8} + \frac{a_h^2}{0.16b^2} - \frac{a_h^4}{0.204b^4}, \tag{4}$$

where $a_h$ is one-half of the width of the acoustical hole. The acoustical holes are modeled by an acoustical resistor $R_h = 12\eta h_b/b^2 L_m^2$ and mass $M_h = 24\rho_0 h_b a_h^2/5b^2 L_m^2$, where $h_b$ is the thickness of the backplate. The acoustical compliance of the backchamber is $C_{\rm bc} = V_{\rm bc}/\rho_0 c^2$, where $V_{\rm bc}$ is the effective volume. A transformer $T_1$ with turn ratio 1: $S_D$ (the area of diaphragm) accounts for the coupling of the acoustical system and mechanical system.

In the mechanical subsystem, the flexural rigidity of the diaphragm is predominantly due to the built-in stress acquired during MEMS fabrication. That renders the compliance of the diaphragm[5] $C_{MD} = 32/\pi^6\sigma_d h_d L_m^2$, where $\sigma_d$ is the built-in stress and $h_d$ is the thickness of the diaphragm. The mechanical mass of the diaphragm is given as $M_{MD} = \rho_d h_d L_m^2$, with $\rho_d$ being the density of the diaphragm material. The mechanical-electrical coupling factor $\phi_T = C_{E0}/C_{EM}$, where $C_{E0}$ is the capacitance due to the dc bias and is given by $C_{E0} = K_{\rm holes}\epsilon_b\epsilon_d\epsilon_0 L_m^2/(h_a\epsilon_b\epsilon_d + h_b\epsilon_d + h_d\epsilon_b)$, and $C_{EM} = (h_a - w_d)/V_{ba}$. Hence, the impedance matrix for the condenser microphone can be written as

$$\mathbf{Z}_{\rm mic} = \begin{bmatrix} Z_E & \dfrac{1}{j\omega C_{EM}} \\ \dfrac{1}{j\omega C_{EM}} & Z_M \end{bmatrix}, \tag{5}$$

where $Z_E$ and $Z_M$ are the equivalent electrical impedance and mechanical impedance, respectively. Note that this matrix is symmetric, as expected for a reciprocal electro-static transducer.

### B. Constant beamwidth design based on the MUSIC template

Assume that $M$ microphones are uniformly distributed with interelement spacing $d$. One benefit of using arrays is the improvement of SNR that can be assessed by the *array white noise gain* (WNG) defined as the increase of the SNR between one sensor and the output of the entire array[2]

$$G = \frac{\rm SNR_{array}}{\rm SNR_{sensor}}.$$

The array inputs ($\mathbf{x}$) and the output ($y$) are given, respectively, by

$$\mathbf{x} = s(t)\mathbf{a} + \mathbf{n},$$

$$y = \mathbf{w}^H\mathbf{x} = s(t)\mathbf{w}^H\mathbf{a} + \mathbf{w}^H\mathbf{n},$$

where $s(t)$ is the source signal, $\mathbf{a} = [1\ \exp[j\omega_c(d\sin\theta/c)]\cdots\exp[j\omega_c(M-1)d\sin\theta]/c]^T$ is called the "manifold vector" associated with the look angle (with respect to the normal) and the center frequency $\omega_c$, $\mathbf{n}$ is the noise vector (assumed to be wide-sense stationary uncorrelated white noise with equal power $\sigma_n^2$), and $\mathbf{w}$ is the array coefficient vector. Hence,

$${\rm SNR_{array}} = \frac{E\{|s(t)\mathbf{w}^H\mathbf{a}|^2\}}{E\{|\mathbf{w}^H\mathbf{n}|^2\}} = \frac{E\{|s(t)|^2\}\mathbf{w}^H\mathbf{a}\mathbf{a}^H\mathbf{w}}{\sigma_n^2\mathbf{w}^H\mathbf{w}}.$$

Recognizing that

$$\frac{E\{|s(t)|^2\}}{\sigma_n^2} = {\rm SNR_{sensor}},$$

we end up with

$$G = \frac{\mathbf{w}^H\mathbf{a}\mathbf{a}^H\mathbf{w}}{\mathbf{w}^H\mathbf{w}}.$$

For example, the array WNG of a benchmark delay-sum array can be calculated by setting $\mathbf{w} = \mathbf{a}/M$:

$$G = \frac{(1/M^2)\mathbf{a}^H\mathbf{a}\mathbf{a}^H\mathbf{a}}{(1/M^2)\mathbf{a}^H\mathbf{a}} = \mathbf{a}^H\mathbf{a} = M.$$

The physical ground to result in this desirable feature is that uncorrelated noises tend to add or cancel with equal probability, while signals add in phase among the array channels.

Our goal is to design a narrow beam throughout a wide range of frequency. Pattern inversion by Fast Fourier Trans-

form (FFT) would not work well in this case because the aperture size of the four-element MEMS microphone array is just too small to produce any meaningful directivity. The spectral leakage of the truncation effect resulting from the small aperture simply destroys the directivity.

To achieve high directivity, a beamforming design based on MUSIC[3] template is adopted. In the MUSIC algorithm, the beam pattern is given as

$$S_{MU}(\theta)=\frac{1}{\mathbf{a}^H\mathbf{P_N}\mathbf{a}}, \tag{6}$$

where the angle $\theta$ and the manifold vector $\mathbf{a}$ are as defined previously, the projection matrix $\mathbf{P}_N=\Sigma_{m=J+1}^{M}\mathbf{u}_m\mathbf{u}_m^H$, $J$ is the number of sources, and $\mathbf{u}_m$ is the $m$th eigenvector of the signal correlation matrix.

The beam pattern produced by the MUSIC is known to be highly directional. However, the MUSIC is mainly intended for finding direction of arrival and it requires intensive computation in the correlation matrix and eigendecomposition. Instead of direct implementation, we adopted a practical approach that utilizes the beam pattern resulting from MUSIC as a design template. That is, we seek the array filters to yield a narrow beam throughout a wide frequency range, as close as possible to that generated by the MUSIC template. To this end, a MUSIC template is generated using Eq. (6) with a source located at 0°, at frequency $\omega_c$, e.g., one-half of the sampling frequency $f_s$. This template with identical subtending angle applies to all frequencies, hence the name ''constant beamwidth.'' The vector of the MUSIC template denoted as $\mathbf{s}_T=[S_{MU}(\theta_1)\cdots S_{MU}(\theta_{Q_d})]^T$ is created by uniformly sampling at $Q_d$ discrete angles ($-90°\leqslant\theta_1\leqslant\cdots\leqslant\theta_{Q_d}\leqslant 90°$) the MUSIC beam pattern in Eq. (6). Next, $P_w$ equally spaced discrete frequencies, $f_r(l)=(f_s/P_w)l$ ($1\leqslant l\leqslant P_w$, are selected within the Nyquist frequency. Let the filter gain at the $m$th element and the $l$th frequency be $w_m(l)$. The array pattern at the $l$th frequency can be written as

$$P_T(l,\theta)=\sum_{m=1}^{M} w_m^*(l)\exp\left(j2\pi f_r(l)\frac{(m-1)d\sin\theta}{c}\right). \tag{7}$$

The symbol * denotes complex conjugate. Since we wish to match an array pattern to the MUSIC template, i.e., $P_T(l,\theta)\approx S_{MU}(\theta)$, the model matching problem for a specific angle can be written as

$$\left[1\ \exp\left(j2\pi f_r(l)\frac{d\sin\theta}{c}\right)\ \cdots\ \exp\left(j2\pi f_r(l)\frac{(M-1)d\sin\theta}{c}\right)\right]\begin{bmatrix} w_1^*(l)\\ w_2^*(l)\\ \vdots\\ w_M^*(l)\end{bmatrix}=S_{MU}(\theta). \tag{8}$$

Assembling all angles, $\theta_1,\ldots,\theta_{Q_d}$ of Eq. (8) leads to the following matrix equation:

$$\begin{bmatrix} 1 & \exp\left(j2\pi f_r(l)\frac{d\sin\theta_1}{c}\right) & \cdots & \exp\left(j2\pi f_r(l)\frac{(M-1)d\sin\theta_1}{c}\right)\\ 1 & \exp\left(j2\pi f_r(l)\frac{d\sin\theta_2}{c}\right) & \cdots & \exp\left(j2\pi f_r(l)\frac{(M-1)d\sin\theta_2}{c}\right)\\ \vdots & \vdots & \ddots & \vdots\\ 1 & \exp\left(j2\pi f_r(l)\frac{d\sin\theta_{Q_d}}{c}\right) & \cdots & \exp\left(j2\pi f_r(l)\frac{(M-1)d\sin\theta_{Q_d}}{c}\right)\end{bmatrix}\begin{bmatrix} w_1^*(l)\\ w_2^*(l)\\ \vdots\\ w_M^*(l)\end{bmatrix}=\begin{bmatrix} S_{MU}(\theta_1)\\ S_{MU}(\theta_2)\\ \vdots\\ S_{MU}(\theta_{Q_d})\end{bmatrix}. \tag{9}$$

Let the array gain vector $\mathbf{w}(l)=[w_1^*(l)\cdots w_M^*(l)]^T$, $\mathbf{s}_T=[S_{MU}(\theta_1)S_{MU}(\theta_2)\ldots S_{MU}(\theta_{Q_d})]^T$, and

$$\mathbf{\Phi}(l)=\begin{bmatrix} 1 & \exp\left(j2\pi f_r(l)\frac{d\sin\theta_1}{c}\right) & \cdots & \exp\left(j2\pi f_r(l)\frac{(M-1)d\sin\theta_1}{c}\right)\\ 1 & \exp\left(j2\pi f_r(l)\frac{d\sin\theta_2}{c}\right) & \cdots & \exp\left(j2\pi f_r(l)\frac{(M-1)d\sin\theta_2}{c}\right)\\ \vdots & \vdots & \ddots & \vdots\\ 1 & \exp\left(j2\pi f_r(l)\frac{d\sin\theta_M}{c}\right) & \cdots & \exp\left(j2\pi f_r(l)\frac{(M-1)d\sin\theta_M}{c}\right)\end{bmatrix}. \tag{10}$$

The optimization problem can be written as a model matching problem

$$\min_{\mathbf{w}(k)}=\|\mathbf{s}_T-\mathbf{\Phi}(l)\mathbf{w}(l)\|_2, \tag{11}$$

where we choose the two norm $\|\cdot\|_2$ as the measure of match. In general, $\mathbf{\Phi}(l)$ is not square and $Q_d>M$. This ''overdetermined'' problem has the least-squares solution, $\mathbf{w}_{\mathrm{LS}}(l)=(\mathbf{\Phi}^H\mathbf{\Phi})^{-1}\mathbf{\Phi}^H\mathbf{s}_T$. Direct calculation of this pseudoinverse

often yields impractical solutions because $\mathbf{\Phi}(l)$ usually has very small singular values. Alternatively, Tikhonov regularization[14] is employed to solve this problem,

$$\mathbf{w}_{\mathrm{LS}}(l)=(\mathbf{\Phi}^H\mathbf{\Phi}+\beta\mathbf{I})^{-1}\mathbf{\Phi}^H\mathbf{s}_T. \tag{12}$$

By varying the parameter $\beta$, we can control the degree of regularization. Repeating the same procedure for $P_w$ frequencies, the frequency response samples of the $m$th channel $w_m^*(l)$ can be obtained. To ensure real impulse responses, the following symmetry of frequency response must be applied

$$\begin{aligned} &H_m(0)=w_m^*(1),\\ &H_m(l)=w_m^*(l), \quad l=1,2,...,P_w,\\ &H_m(l)=w_m(l), \quad l=(P_w+1),(P_w+2),\ldots,(2P_w-1). \end{aligned} \tag{13}$$

The impulse response for each channel is then calculated with the aid of inverse FFT,

$$\begin{aligned} &h_m(k)=\frac{1}{2P_w}\sum_{l=0}^{2P_w-1}H_m(l)e^{j(\pi/P_w)lk},\\ &k=0,1,...,(2P_w-1). \end{aligned} \tag{14}$$

The resulting $h_m(k)$ is generally noncausal and circular shift is required to obtain a causal impulse response $h'_m(k)$.

Since we wished to simplify the array design on the chip, we chose simple implementation using analog filters. The frequency responses of the FIR filters $h'_m(k)$ are calculated

$$H'_m(l)=\sum_{k=0}^{2P_w-1}h'_m(k)e^{-j(\pi/P_w)lk}, \quad l=0,1,...,(2P_w-1). \tag{15}$$

The transfer functions of the analog filters are fitted with the MATLAB[15] command *invfreqs*. The resulting analog filters of array beamforming are realized by using active filter techniques such as the biquad circuits.[16]

### C. Discussion of self-noise

The self-noise, or background noise, is defined as the *rms* output signal of the microphone in the absence of acoustical excitation. An in-depth comparison of the background noise in piezoresistive, eletret condenser, and ceramic microphones can be found in the paper by Zuckerwar *et al.*[17] Their analysis identified four types of background sources:

(1) mechanical Johnson noise due to the Brownian motion of air molecules impinging on the diaphragm, as represented by a damping resistor $R_A$, and $1/f$ (inversely proportional to frequency) noise generated by motion of the diaphragm;
(2) thermal channel noise generated in the FET channel;
(3) gate shot noise generated in the FET;
(4) electrical Johnson noise generated in the drain resistor $R_d$.

An equivalent circuit analysis in their work led to the following expression of the power spectral density of background noise in terms of the above-mentioned sources:

TABLE I. The parameters assumed in the FDM procedure for the quasistatic analysis of the MEMS condenser microphone.

| Diaphragm material | Silicon nitride |
|---|---|
| Diaphragm length ($L_m$) | 1 mm |
| Diaphragm thickness ($h_d$) | 1 $\mu$m |
| Air gap distance ($h_a$) | 3 $\mu$m |
| Acoustical holes number ($N_m$) | 100 |
| Acoustical holes side length ($a_d$) | 60 $\mu$m |
| Vacuum permittivity ($e_0$) | $8.85\times10^{-12}$ F/m |
| Dielectric constant of the diaphragm ($e_d$) | 7 |
| Young's modulus ($E$) | 300 GPa |
| Built-in stress ($\sigma_d$) | 50 MPa |
| Poisson's ratio ($\upsilon$) | 0.2 |
| Bias voltage ($V_{ba}$) | 50 V |
| Pressure loading ($P_{sp}$) | 5 Pa |

$$J=\frac{S^2(4kTR_A+A_m^2/f)}{(1-\omega^2/\omega_0^2)^2+(\omega/\omega_0Q)^2}+\left[\frac{4(0.65)kT}{g_m}+\frac{2qI_G}{(\omega C_{t0})^2}\right](g_mR_d)^2+4kTR_d. \tag{16}$$

In Eq. (16), $S$ is the diaphragm area, $k$ is the Boltzmann constant, $T$ is the absolute temperature, $f$ and $\omega$ are frequencies in Hz and rad/s, respectively, $\omega_0$ and $Q$ are the resonance and quality factor of the diaphragm, $A_m^2$ is the $1/f$ coefficient, $g_m$ is the FET transconductance, $C_{t0}$ is the diaphragm-backplate capacitance, $q$ is the electronic charge, and $I_G$ is the gate leakage current. As revealed in Eq. (16), the gate shot noise introduces a $1/\omega^2$ component. At frequencies well below the resonance, the membrane damping (Brownian) component behaves as a Johnson noise source, which dominates over the channel and drain resistor components.

An important finding in their work is a formula obtained by linear regression that correlates the $1/f$ noise coefficient ($b_1$) with the diaphragm damping resistor $R_A$.

$$\log(b_1)=-24.1+1.89\log(R_A). \tag{17}$$

This empirical relationship applies to any diaphragm-based microphone, independent of detection technology. Because the diaphragm damping resistance increases inversely with diaphragm diameter, one may expect a significant $1/f$ component in MEMS microphones. As a caveat, the SNR of a microphone with overly damped diaphragm can be quite poor at low frequencies due to the associated $1/f$ noise.

## III. NUMERICAL AND EXPERIMENTAL INVESTIGATIONS

### A. Quasistatic analysis

Following the aforementioned iterative procedure of FDM, the quasistatic analysis was carried out to predict the collapse condition. Table I shows the parameters assumed in this simulation. The diaphragm material was assumed to be silicon nitride. The FDM calculation was based on a mesh with $25\times25$ grid points, covering a quarter of the diaphragm. The built-in stress=50 MPa and the Young's modulus=300 GPa is assumed for a 1 mm square diaphragm made of sili-

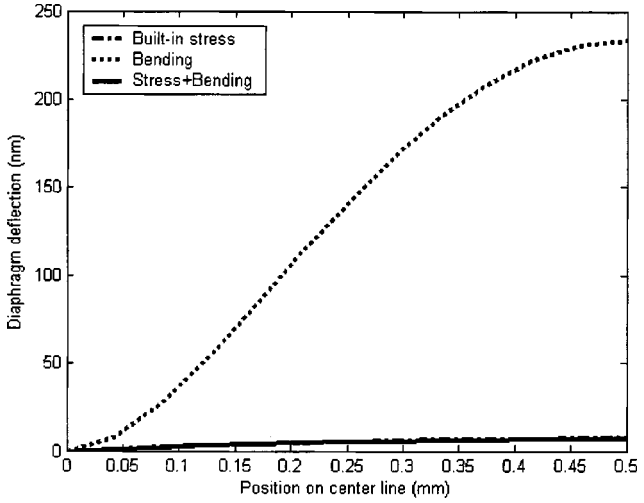


FIG. 3. The deflection profiles calculated by the FDM for three different conditions: bending effect only, built-in stress only, and "stress+bending." The diaphragm material was assumed to be silicon nitride. The built-in stress=50 MPa, the Young's modulus=300 GPa, and the diaphragm is subject to a uniform load of 5 Pa.

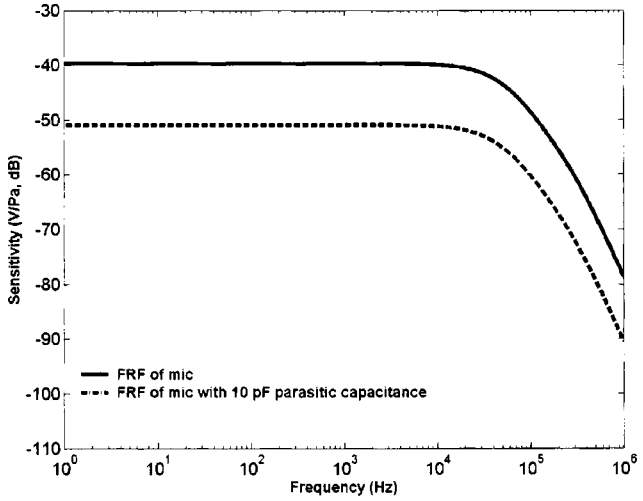


FIG. 4. The frequency response of the MEMS condenser microphone simulated by using the PSPICE. The effects of parasitic capacitance on sensitivity are also indicated in the same plot. In the simulation, 10 pF parasitic capacitance is assumed.

con nitride.[18] The calculated deflection profiles for three different conditions including "bending effect only," "built-in stress only," and "stress+bending" are shown in Fig. 3. The results indicate that the effect due to built-in stress is the dominant factor in overall diaphragm stiffness (about two orders higher than the bending effect). If only the built-in stress is considered, the maximum deflection at the center of the diaphragm reaches 7 nm, which is of the same order as the value 3.3 nm predicted by a parallel plate/linear spring model with the spring constant deduced from $C_{MD}=32/\pi^6\sigma_d h_d L_m^2$. Using this FDM model, the critical bias voltage near collapse was found to be 84 V for the present setting. For stability of the device, a conservative value, 60% of the previously calculated critical voltage, or 50 V, was thus selected to be the dc bias. The maximum static deflection of the diaphragm corresponding to such dc bias voltage attained 0.32 $\mu$m. Since the predicted deflection is far less than the nominal air gap (3 $\mu$m), the condenser microphone should be rather safe during operation without the risk of collapse.

TABLE II. The parameters assumed in the linear dynamic analysis of the MEMS condenser microphone.

| Diaphragm material | Silicon nitride |
|---|---|
| Diaphragm length ($L_m$) | 1 mm |
| Diaphragm thickness ($h_d$) | 1 $\mu$m |
| Diaphragm density ($\rho_d$) | 3440 kg/m$^3$ |
| Backplate thickness ($h_b$) | 20 $\mu$m |
| Air gap distance ($h_a$) | 3 $\mu$m |
| Air density ($\rho_0$) | 1.3 kg/m$^3$ |
| Sound speed ($c$) | 343 m/s |
| Dynamic viscosity of air ($\eta$) | $1.86\times10^{-5}$ N s/m$^2$ |
| Acoustical holes number ($N_m$) | 100 |
| Acoustical holes side length ($a_d$) | 60 $\mu$m |
| Vacuum permittivity ($e_0$) | $8.85\times10^{-12}$ F/m |
| Dielectric constant of the diaphragm ($e_d$) | 2 |
| Dielectric constant of the back plate ($e_b$) | 2 |
| Young's modulus ($E$) | 300 GPa |
| Built-in stress ($\sigma_d$) | 50 MPa |
| Poisson's ratio ($\upsilon$) | 0.2 |
| Bias voltage ($V_{ba}$) | 50 V |
| Pressure loading ($P_{sp}$) | 5 Pa |

### B. Simulation of linear dynamic model and preamplifier circuits

The linear dynamic model is simulated with the aid of the PSPICE.[19] Table II summarizes the parameters assumed in this simulation. For simplicity, the linear analysis considers only the dynamic response, but not the dc deflection of the diaphragm caused by the bias voltage. The partial pull-down may change the effective compliance, increasing it or decreasing it depending on the proximity to snap-down. Increasing (decreasing) the compliance reduces the bandwidth (responsivity). Figure 4 shows the simulated frequency response between the incident sound pressure and the open-circuit output voltage of the microphone with a 1 mm$\times$1 mm

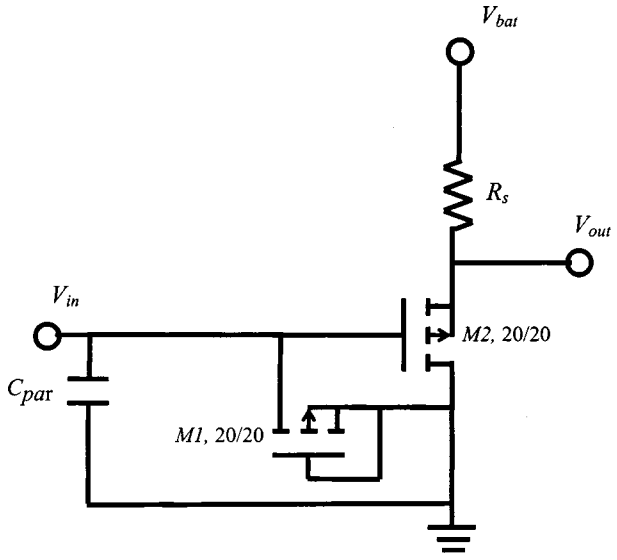


FIG. 5. The PMOS source-follower preamplifier. Two PMOS transistors of 20 $\mu$m/20 $\mu$m are selected to implement the amplifier.

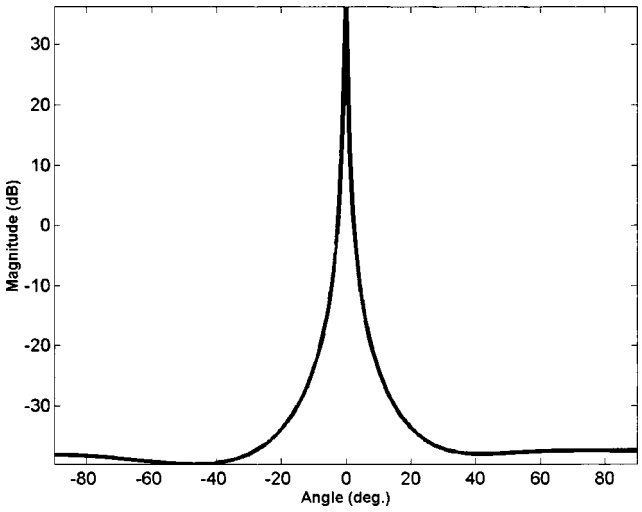


FIG. 6. A template obtained using MUSIC algorithm for $Q_d=181$, $P_w=256$. The beam pattern of this template exhibits a sharp peak at 0° with beamwidth approximately 10° for 40 dB attenuation.

square diaphragm. The 3 dB cut-off frequency occurs at 42.7 kHz. The frequency response is relatively flat throughout the band 0–42 kHz. The predicted open-circuit sensitivity was approximately −39.7 dB V/Pa, or 10.35 mV/Pa. The obtained sensitivity of MEMS microphone is comparable to conventional 1/4 in. condenser microphones.

The preamplifier shown in Fig. 5 is needed because of the relatively large output impedance of the condenser microphone. This buffering preamplifier utilizes the PMOS transistor as the bias element, owing to its high impedance characteristic. The optimum SNR is achieved by having much higher input impedance than the sensor impedance. After the preamplifier, the signal should be robust and can be taken off chip if necessary. However, on-chip integration of the buffering preamplifier is still important to reduce parasitics. Figure 4 also plots the microphone frequency response if the parasitic capacitance $C_{\text{par}}=10\,\text{pF}$ is present. The loss of sensitivity is frequency dependent. The sensitivity of MEMS

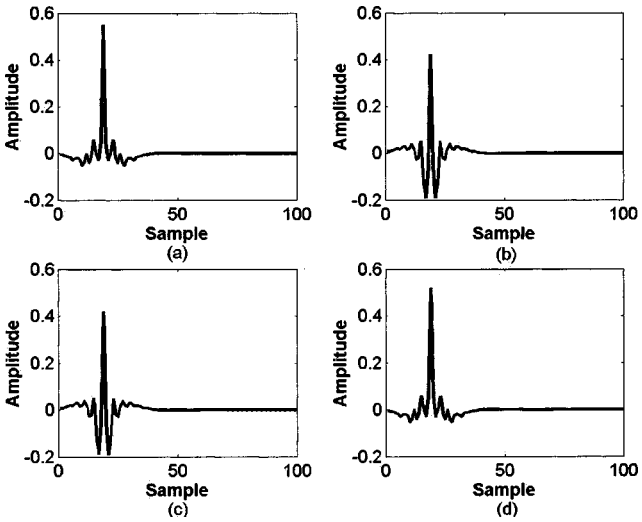


FIG. 7. The calculated impulse responses of the array FIR filters after the circular shift. (a) Channel one. (b) Channel two. (c) Channel three. (d) Channel four.

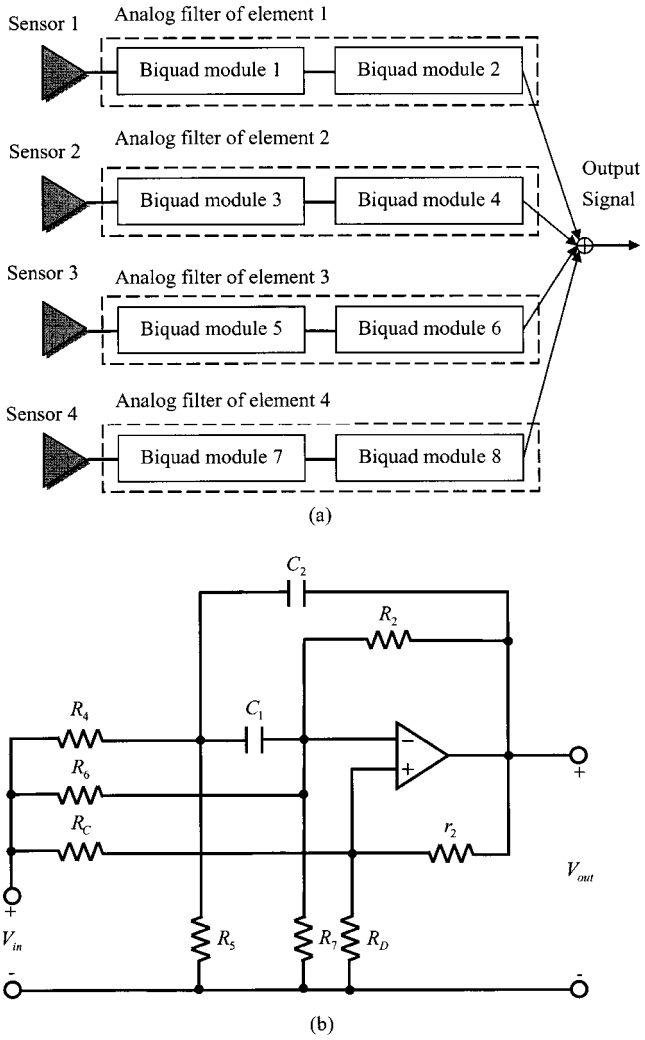


FIG. 8. Analog implementation of the array filters. (a) Four channel microphone array implemented by the biquad modules. (b) Detailed biquad circuit.

microphones is significantly reduced by 10 dB up to the cutoff frequency due to the parasitic capacitance. As evident from this result, the SoC design is crucial for MEMS microphones to yield sufficient sensitivity.

### C. Constant beamwidth design based on the MUSIC template

In this section, the constant beamwidth design based on the MUSIC template is presented. The array contains four microphones in a line, with interelement spacing 5 mm. A MUSIC template is constructed using Eq. (6) for $Q_d=181$ discrete angles and $P_w=256$ discrete frequencies. The beam pattern of the resulting template is shown in Fig. 6. This is the MUSIC spectrum plotted in the angle space, and not the conventional Fourier spectrum. This template exhibits a sharp peak at 0° with beamwidth approximately 10° for 40 dB attenuation. After obtaining the frequency response of the microphone array, the impulse response, or the FIR filter coefficients, are calculated using inverse FFT. Circular shift is applied to obtain causal FIR filters. The resulting impulse responses of the array FIR filters are shown in Fig. 7. The MATLAB function *invfreqs* is used to find the continuous time

TABLE III. The parameters used in the analog biquad circuits of the array filters.

| | Module 1 | Module 2 | Module 3 | Module 4 |
|---|---|---|---|---|
| $R_2$ | $2.65\times10^3\ \Omega$ | $3.64\times10^3\ \Omega$ | $1.75\times10^3\ \Omega$ | $5.60\times10^3\ \Omega$ |
| $R_4$ | $8.31\times10^2\ \Omega$ | $7.68\times10^3\ \Omega$ | $7.89\times10^2\ \Omega$ | $1.07\times10^4\ \Omega$ |
| $R_5$ | Infinite | $3.44\times10^3\ \Omega$ | Infinite | $3.18\times10^3\ \Omega$ |
| $R_6$ | Infinite | $4.34\times10^3\ \Omega$ | Infinite | $2.83\times10^4\ \Omega$ |
| $R_7$ | $7.87\times10^2\ \Omega$ | $2.22\times10^4\ \Omega$ | $2.75\times10^2\ \Omega$ | $1.74\times10^4\ \Omega$ |
| $R_C$ | $6.82\times10^3\ \Omega$ | $2.39\times10^3\ \Omega$ | $9.75\times10^3\ \Omega$ | $5.25\times10^3\ \Omega$ |
| $R_D$ | $1.17\times10^3\ \Omega$ | $1.72\times10^3\ \Omega$ | $1.11\times10^3\ \Omega$ | $1.24\times10^3\ \Omega$ |
| $r_2$ | $10.0\times10^3\ \Omega$ | $10.0\times10^3\ \Omega$ | $10.0\times10^3\ \Omega$ | $10.0\times10^3\ \Omega$ |
| $C_1$ | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F |
| $C_2$ | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F |
| | Module 5 | Module 6 | Module 7 | Module 8 |
| $R_2$ | $2.26\times10^3\ \Omega$ | $5.52\times10^3\ \Omega$ | $1.64\times10^3\ \Omega$ | $3.82\times10^3\ \Omega$ |
| $R_4$ | $7.88\times10^2\ \Omega$ | $8.59\times10^3\ \Omega$ | $8.15\times10^2\ \Omega$ | $1.32\times10^4\ \Omega$ |
| $R_5$ | Infinite | $3.42\times10^3\ \Omega$ | Infinite | $2.78\times10^3\ \Omega$ |
| $R_6$ | Infinite | $2.19\times10^4\ \Omega$ | Infinite | $7.71\times10^3\ \Omega$ |
| $R_7$ | $4.26\times10^2\ \Omega$ | $1.99\times10^4\ \Omega$ | $2.76\times10^2\ \Omega$ | $7.18\times10^4\ \Omega$ |
| $R_C$ | $6.83\times10^3\ \Omega$ | $4.20\times10^3\ \Omega$ | $1.63\times10^4\ \Omega$ | $4.15\times10^3\ \Omega$ |
| $R_D$ | $1.17\times10^3\ \Omega$ | $1.31\times10^3\ \Omega$ | $1.07\times10^3\ \Omega$ | $1.32\times10^3\ \Omega$ |
| $r_2$ | $10.0\times10^3\ \Omega$ | $10.0\times10^3\ \Omega$ | $10.0\times10^3\ \Omega$ | $10.0\times10^3\ \Omega$ |
| $C_1$ | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F |
| $C_2$ | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F | 0.01 $\mu$F |

transfer functions of the array filters, applicable to the frequency range 0–8 kHz. Next, these array filters are implemented by using analog biquad circuits. Since each analog transfer function is of order four, two biquad circuits shown in Fig. 8 are cascaded to realize the array filter. The parameters of the biquad circuits are shown in Table III. The frequency responses obtained using the simulation and the experiment, respectively, are compared in Fig. 9. The equivalent filter weights at four frequencies are shown in Table IV. The magnitude and phase responses are in good agreement. It is worth noting that the 180° phase inversion at 6 kHz for elements 2 and 3 results in the increased directivity.

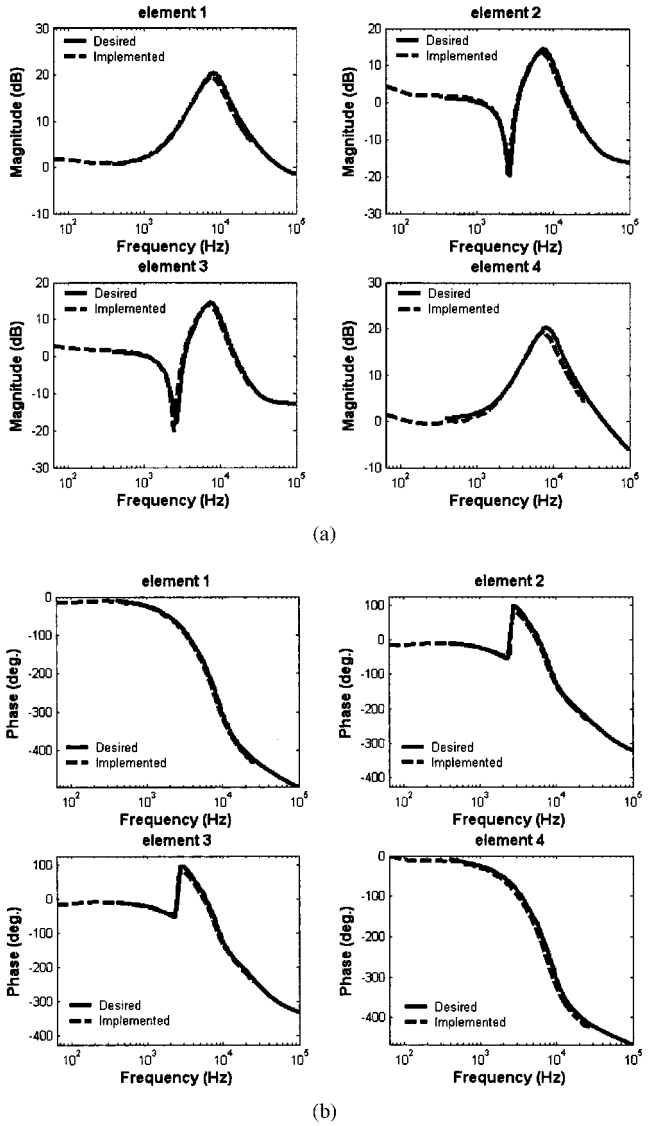


FIG. 9. Comparison of the frequency responses of the analog array filters obtained using the simulation and the experiment, respectively. (a) Magnitude response. (b) Phase response.

### D. Experimental investigation

In order to verify the proposed array microphone system, experiments were carried out using a commercial product of silicon condenser microphone, SP0103NC3-2 of EMKAY. This omni-directional microphone has the frequency range of 100 Hz–10 kHz and sensitivity of −42 dB V/Pa at 1 kHz. With the same configuration as in the simulation, a linear array comprising four elements of such microphones with 5 mm interelement spacing was constructed, as shown in Fig. 10. The previously designed array filters for these four microphone elements were realized by using operational amplifier circuit. Using this experimental arrangement, tests were conducted inside an anechoic chamber to evaluate the SNR as well as the directivity for various microphone configurations.

First experiment compares the SNRs of one single microphone and the four-element array. In each case, the SNR is calculated by subtracting the power of the sensor output when the microphones are exposed to a white noise input band-limited to 20 kHz, from the power of the sensor output when the source is switched off. Hence, the noise would be the combination of the electronic self-noise and the background noise inside the anechoic chamber. The thus measured SNRs were found to be 24.1 and 35.5 dB for the single microphone and the array microphone, respectively. The SNR was improved by 11.4 dB, which was not far from the theoretical prediction, 12 dB. This experimental result suggests that array configuration is capable of enhancing SNR, as compared to one sensor element. However, the broadband measurement does not assess the performance at 6 kHz.

A second experiment compares the measured directional responses of the silicon microphones for three cases. The results are shown in Fig. 11. The first case labeled in Fig. 11(a) as ''one microphone'' refers to the measured directivity of a single microphone. The second case labeled in Fig. 11(b) as ''four microphones'' refers to the measured directivity produced by the direct sum of four microphone outputs. The third case labeled in Fig. 11(c) as ''four microphones with filters'' refers to the measured directivity produced by

TABLE IV. The equivalent filter weights and directivity indices at four frequencies.

| Frequency (Hz) | Element | Weights: Magnitude (dB) | Weights: Phase (degree) | Directivity index (dB): Delay Sum | Directivity index (dB): MUSIC |
|---|---|---|---|---|---|
| | 1 | 0.678 36 | −15.87 | | |
| 500 | 2 | 1.586 | −13.619 | 0.004 | 0.004 |
| | 3 | 1.352 | −13.576 | | |
| | 4 | 0.475 28 | −15.285 | | |
| | 1 | 2.498 5 | −27.868 | | |
| 1000 | 2 | 0.602 49 | −24.254 | 0.015 | 0.016 |
| | 3 | 0.360 99 | −24.558 | | |
| | 4 | 1.260 2 | −31.171 | | |
| | 1 | 5.659 4 | −52.9 | | |
| 2000 | 2 | −4.723 8 | −44.433 | 0.060 | 0.087 |
| | 3 | −5.925 4 | −43.677 | | |
| | 4 | 4.780 6 | −58.882 | | |
| | 1 | 17.578 | −194.06 | | |
| 6000 | 2 | 12.482 | −14.345 | 0.533 | 1.997 |
| | 3 | 12.652 | −14.392 | | |
| | 4 | 17.859 | −202.77 | | |

filtering the four microphone outputs with the aforementioned analog filter circuits. Comparison of the experimental results of cases 1 and 2 reveals that only minor improvement on directivity is obtained if the array outputs are directly summed. However, significant improvement on directivity is observed by comparing cases 1 and 3 if the array outputs are preprocessed by the previously designed constant beamwidth array filters. Specifically, at the frequency 6 kHz, deep notches appear at the end-fire direction and the 10 dB beamwidth reaches approximately $\pm 60°$ when the four microphone outputs are filtered with the aforementioned analog filter circuits. This is also revealed in the measured directivity indices at four frequencies in Table IV: 0.533 dB for the delay-sum array versus 1.997 for the MUSIC array at 6 kHz.

## IV. CONCLUSIONS AND FUTURE PERSPECTIVES

The present work represents a multidisciplinary effort that combines the knowledge in electroacoustic transducers, array signal processing, and MEMS technology. The original contribution of this paper can be summarized as follows. First, by taking advantage of MEMS process, we have ac-

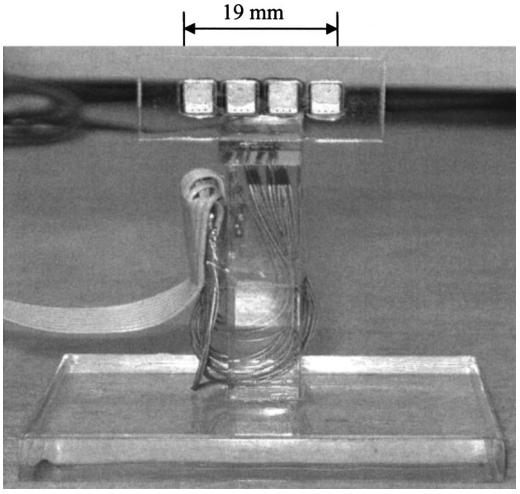


FIG. 10. Experimental arrangement of the four-element MEMS microphone array.

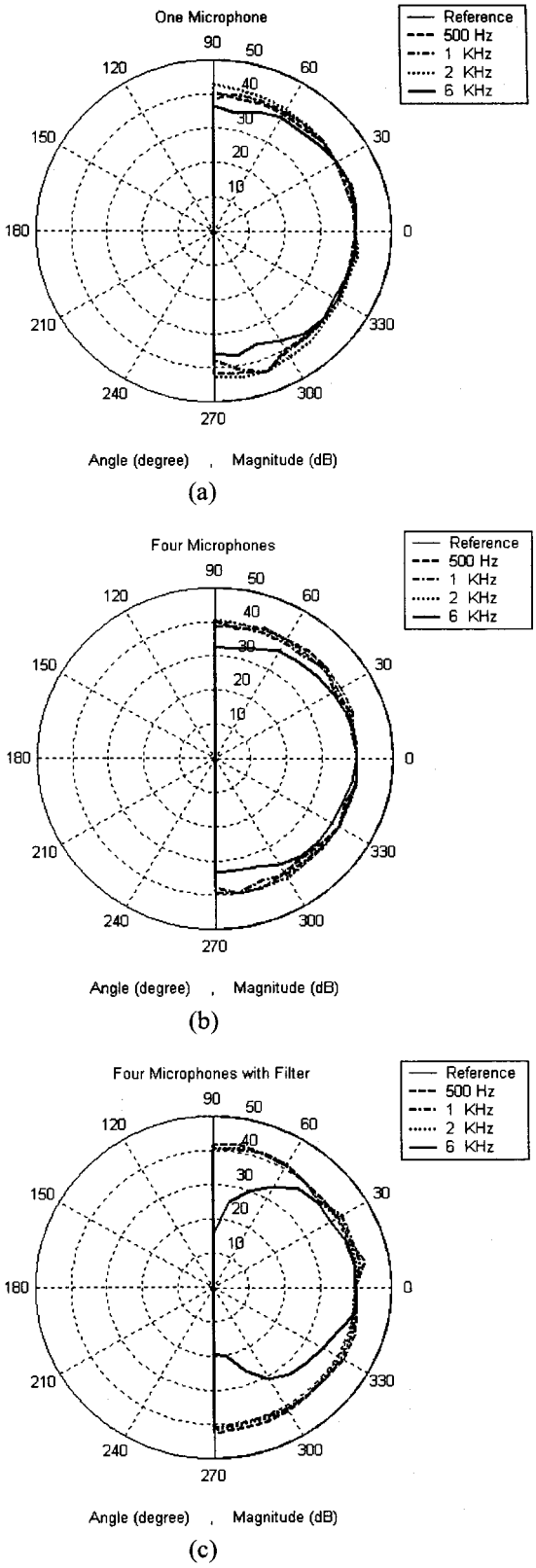


FIG. 11. Comparison of the measured directional responses of the silicon microphones. (a) The measured directivity of a single microphone. (b) The measured directivity produced by the direct sum of four microphone outputs. (c) The measured directivity produced by filtering the four microphone outputs with the analog filter circuits.

complished a system design which attempts to integrate a microphone array and the associated circuit on one chip. This differs from common approaches of microphone arrays that are predominantly based on single channel conventional microphones. Second, a beamforming technique more sophisticated than simple delay-sum method was developed in this study to improve the SNR and directionality. A least-squares model match procedure is proposed in the synthesis of the array filters to modify a DOA algorithm into a practical beamforming method. Third, in addition to the array beamforming design, the proposed idea was actually implemented and verified through experiments. The array filters are based on analog implementation that avoids the complexity of digital implementation.

Apart from the work accomplished in this study, certain limitations concerning the proposed system should be mentioned. Although the SNR is improved using the array configuration, little directivity is seen in the results of the array for frequencies below 5 kHz, which is the range for human speech. To further enhance the array directivity, a number of solutions are planned for the future research. The array configuration adopted in the present study is the so-called ''dimensional array.''[20] Directivity of such arrays is frequency-dependent and proportional to the aperture size of the array. Under the physical constraint of the small aperture of the present MEMS microphone, instead of the present analog implementation, high directivity can possibly be attained by using direct digital implementation of the beamforming algorithms. Yet, another possibility is using the so-called differential array.[20] In contrast to the dimensional arrays, the directivity of differential arrays is independent of frequency and aperture size. However, the price to pay using this approach is that its frequency response has $\omega^n$ ($\omega$: frequency, $n$: element number $-1$) dependence. This is a characteristic with differential arrays—higher directivity but lower array WNG, and hence lower dynamic range. This suggests that equalization is required to boost the gain at the low frequencies, while suppressing the noise amplification at the high frequencies. This could be quite challenging for a MEMS device since the SNR is generally a crucial problem at low frequencies due to the previously mentioned $1/f$ noise component. The issue of whether both attributes—SNR improvement and good directivity in terms of array configurations—can be best compromised will be addressed in future research.

## ACKNOWLEDGMENTS

The work was supported by the Nation Science Council in Taiwan, Republic of China, under Project No. NSC 91-2212-E009-032. The authors also wish to thank Knowles Electronics, Inc. for free samples of the MEMS microphones.

[1] M. Pedersen, ''A polymer condenser microphone realized on silicon containing preprocessed integrated circuits,'' Ph.D. dissertation, University of Twente, The Netherlands, 1997.
[2] H. L. Van Trees, *Optimal Array Processing* (Wiley, New York, 2002).
[3] R. O. Schmidt, ''Multiple emitter location and signal parameter estimation,'' IEEE Trans. Antennas Propag. **AP-34**, 276–280 (1986).
[4] D. Hohm and G. Hess, ''A subminiature condenser microphone with silicon nitride membrane and silicon back plate,'' J. Acoust. Soc. Am. **85**, 476–480 (1989).
[5] W. Kühnel and G. Hess, ''A silicon condenser microphone with structured back plate and silicon nitride membrane,'' Sens. Actuators, A **30**, 251–258 (1992).
[6] J. Bergqvist, ''Finite element modeling and characterization of a silicon condenser microphone with a highly perforated back plate,'' Sens. Actuators, A **39**, 191–200 (1993).
[7] H. F. Olson, *Acoustical Engineering* (Van Nostrand, Princeton, NJ, 1957).
[8] L. L. Beranek, *Acoustics* (McGraw–Hill, New York, 1954).
[9] M. Pedersen, W. Olthuis, and P. Bergveld, ''High-performance condenser microphone with fully integrated CMOS amplifier and DC-DC voltage converter,'' J. Microelectromech. Syst. **7**, 387–394 (1998).
[10] S. Chowdhury, M. Ahmadi, G. A. Jullien, and W. C. Miller, ''A MEMS Implementation of an Acoustical Sensor Array,'' Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS 2001), Sydney, Australia, 2 May, 2001, pp. 273–276.
[11] D. P. Arnold, T. Nishida, L. N. Cattafesta, and M. Sheplak, ''A directional acoustic array using silicon micromachined piezoresistive microphones,'' J. Acoust. Soc. Am. **113**, 289–298 (2003).
[12] S. P. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells* (McGraw–Hill, New York, 1959).
[13] Z. Skvor, ''On the acoustical resistance due to viscous losses in the air gap of electrostatic transducers,'' Acoustica **19**, 295–299 (1968).
[14] A. Schuhmacher, J. Hald, K. B. Rasmussen, and P. C. Hansen, ''Sound source reconstruction using inverse boundary element calculations,'' J. Acoust. Soc. Am. **113**, 114–127 (2003).
[15] A. Grace, A. J. Laub, J. N. Little, and C. M. Thompson, *Matlab Control System Toolbox* (The Math Works, Natick, MA, 1999).
[16] W. K. Chen, *Passive and Active Filters: Theory and Implementations* (Wiley, New York, 1986).
[17] A. J. Zuckerwar, T. R. Kuhn, and R. M. Serbyn, ''Background noise in piezoresistive, electret condenser, and ceramic microphones,'' J. Acoust. Soc. Am. **113**, 3179–3187 (2003).
[18] M. Madou, *Fundamentals of Microfabrication* (CRC Press, New York, 1997).
[19] *OrCad, PSPICE, Mixed A/D Circuit Simulator* (Cadence Design Systems, San Jose, CA, 2002).
[20] G. M. Sessler and J. H. West, ''Directional Transducers,'' IEEE Trans. Audio Electroacoust. **AU-19**, 19–23 (1971).