

# UC San Diego

## UC San Diego Previously Published Works

### Title

Micropower Mixed-Signal VLSI Independent Component Analysis for Gradient Flow Acoustic Source Separation

### Permalink

<https://escholarship.org/uc/item/23k3z8tp>

### Journal

IEEE Transactions on Circuits and Systems I Regular Papers, 63(7)

### ISSN

1057-7130

### Authors

Stanaćević, Milutin

Li, Shuo

Cauwenberghs, Gert

### Publication Date

2016-07-01

### DOI

10.1109/tcsi.2016.2556122

Peer reviewed



Published in final edited form as:

*IEEE Trans Circuits Syst I Regul Pap.* 2016 July ; 63(7): 972–981. doi:10.1109/TCSI.2016.2556122.

## Micropower Mixed-signal VLSI Independent Component Analysis for Gradient Flow Acoustic Source Separation

**Milutin Stana evi [Member, IEEE],**

Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794-2350

**Shuo Li [Student Member, IEEE], and**

Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794-2350

**Gert Cauwenberghs [Fellow, IEEE]**

Department of Bioengineering, University of California San Diego, La Jolla, CA 92093

### Abstract

A parallel micro-power mixed-signal VLSI implementation of independent component analysis (ICA) with reconfigurable outer-product learning rules is presented. With the gradient sensing of the acoustic field over a miniature microphone array as a pre-processing method, the proposed ICA implementation can separate and localize up to 3 sources in mild reverberant environment. The ICA processor is implemented in 0.5  $\mu\text{m}$  CMOS technology and occupies 3 mm  $\times$  3 mm area. At 16 kHz sampling rate, ASIC consumes 195  $\mu\text{W}$  power from a 3 V supply. The outer-product implementation of natural gradient and Herault-Jutten ICA update rules demonstrates comparable performance to benchmark FastICA algorithm in ideal conditions and more robust performance in noisy and reverberant environment. Experiments demonstrate perceptually clear separation and precise localization over wide range of separation angles of two speech sources presented through speakers positioned at 1.5 m from the array on a conference room table. The presented ASIC leads to a extreme small form factor and low power consumption microsystem for source separation and localization required in applications like intelligent hearing aids and wireless distributed acoustic sensor arrays.

### Index Terms

Blind source separation; Independent component analysis; Micropower techniques

## I. Introduction

Blind source separation(BSS) has long been considered a hard signal processing problem and different algorithms for a wide range of applications in speech processing [1], wireless communications [2] and biomedical signal processing [3] exist. Independent component analysis (ICA) is a signal processing technique for solving instantaneous BSS problem that

can be formulated as a linear transformation that minimizes the statistical dependence between components in a random data vector [4]. Few analog VLSI implementations of ICA exist in the literature [5], [6], while in the digital domain, the high-power implementations using FPGAs are common practice in the field [7], [8] and few digital ASICs [9], [10] are reported.

For blind separation of acoustic sources new opportunities arise with the miniaturization of microphones, specially with the design of MEMS microphones that can capture the field distribution of the impinging sound source [11], [12]. While the human auditory system performs remarkably well in segregating multiple streams of acoustic sources, performance of modern hearing aids significantly deteriorates in the presence of multiple signal and noise sources in the acoustic scene [13]. To be effective in resolving the signal of interest, both localization and separation of multiple acoustic sources are required. We have demonstrated that the direction of acoustic wave propagation can be estimated by differential spatial sensing of the field on sub-wavelength scale, using gradient flow technique [14], [15]. Mixed-signal VLSI implementation of the method [16] has demonstrated improved performance in terms of power dissipation and bearing resolution over conventional bearing estimation localizers. Besides its use in bearing estimation, gradient flow provides an efficient signal representation as a front-end for blind source separation of traveling wave signals. In the presence of multiple signal sources, gradient flow converts the problem of separating unknown delayed mixtures of independent signal sources, into a simpler problem of separating instantaneous mixtures of the time-differentiated source signals [14]. This formulation is equivalent to the problem statement in independent component analysis. Implementation of ICA in mixed-signal VLSI combined with spatial gradient sensing ASIC yields a real-time low-power system for separation and localization of multiple independent signal sources impinging on miniature microphone array. We have previously reported the implementation of the ICA update rule formulated in the outer-product form with the fixed-diagonal terms [17], [18]. The presented ICA architecture implements natural gradient update rule with the back-propagation path providing improved stability and faster convergence.

The paper is organized as follows. Section II reviews algorithms for static linear ICA and describes how gradient flow yields to joint localization and separation in case of multiple travelling source waves. A general mixed-signal parallel architecture, that can be configured for implementation of various ICA update rules is presented in Section III. Experimental characterization of a fabricated prototype and demonstration of the separation performance is presented in Section IV, followed by the concluding comments in Section V.

## II. Blind Source Separation

The blind source separation problem can be formulated in a following manner.  $N$  unknown sources  $\mathbf{s}(t) = [s_1(t) s_2(t) \dots s_N(t)]^T$  propagate through an unknown medium and are observed by an array of  $M$  sensors  $\mathbf{x}(t) = [x_1(t) x_2(t) \dots x_M(t)]^T$ . The task is to obtain estimates of the source signals  $\mathbf{y}(t) = [y_1(t) y_2(t) \dots y_N(t)]^T$  from the observed sensor signals. The common assumption on the source signals utilized in solving the separation problem is that the source signals are statistically independent [4]. The model of the mixing process is determined in accordance

with the setting of the problem. In the case of linear instantaneous mixing, the observed signals  $\mathbf{x}$  are linear combinations of the unknown source signals

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

where  $\mathbf{A}$  is  $M \times N$  dimensional mixing matrix and  $\mathbf{n}$  is the additive noise. When the number of source  $N$  is greater than the number of observations  $M$ ,  $N > M$ , some prior information on the source signals is necessary for solving the problem.

### A. Gradient Flow Signal Representation

The linear instantaneous mixing model is not applicable when the source signals are traveling waves like acoustic waves. We devised a signal conditioning technique gradient flow [14] for unmixing of the observed signals on the miniature microphone array with the distance between microphones much smaller than the wavelength of the source signals. Gradient flow separates and localizes sources by relating spatial and temporal derivatives of the impinging acoustic field, as illustrated in Figure 1(a). Observation of the first order spatial gradients of the acoustic field  $\xi_{10}(t)$  and  $\xi_{01}(t)$  in perpendicular directions  $p$  and  $q$  in the plane in the case of a single source enables indirect estimation of time delays  $\tau_p$  and  $\tau_q$ , as illustrated in Figure 1(b),

$$\xi_{10}(t) \approx \tau_p \dot{\xi}_{00}(t)$$

$$\xi_{01}(t) \approx \tau_q \dot{\xi}_{00}(t), \quad (2)$$

where  $\dot{\xi}_{00}(t)$  is the time-differentiated spatial common mode of the acoustic field  $\xi_{00}(t)$ . The propagation delays  $\tau_p$  and  $\tau_q$  directly yield the azimuth  $\theta$  and the elevation  $\phi$  angle of the source.

When multiple sources are impinging on the miniature array, as illustrated in Figure 1(c) for three sources, the observed first order spatial gradients yield linearly mixed observations of the time-differentiated source signals  $\dot{s}_l$  each scaled by propagation delays  $\tau_p^l$  and  $\tau_q^l$  along the gradient directions:

$$\xi_{10}(t) \approx \sum_l \tau_p^l \dot{s}_l(t) + \nu_{10}(t)$$

$$\xi_{01}(t) \approx \sum_l \tau_q^l \dot{s}_l(t) + \nu_{01}(t), \quad (3)$$

where  $\nu_{10}$  and  $\nu_{01}$  represent spatial derivative components of additive noise in the sensor observations.

Likewise, the time-differentiated spatial common mode  $\dot{\xi}_{00}(t)$  yields further linearly mixed observation of the time-differentiated source signals  $\dot{s}_l$

$$\dot{\xi}_{00}(t) \approx \sum_l \dot{s}_l(t) + \dot{\nu}_{00}, \quad (4)$$

where  $\dot{\nu}_{00}$  represents the time-derivative of the common mode component of additive noise in the sensor observations. Looking at gradient signals  $\dot{\xi}_{00}(t)$ ,  $\xi_{10}(t)$  and  $\xi_{01}(t)$ , equations (3) and (4) can be identified as a linear instantaneous mixture of time-differentiated source signals, which is in the form of classic linear static ICA (1) as

$$\begin{bmatrix} \dot{\xi}_{00} \\ \xi_{10} \\ \xi_{01} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 1 \\ \tau_p^1 & \cdots & \tau_p^N \\ \tau_q^1 & \cdots & \tau_q^N \end{bmatrix} \begin{bmatrix} \dot{s}_1 \\ \vdots \\ \dot{s}_N \end{bmatrix} + \begin{bmatrix} \dot{\nu}_{00} \\ \nu_{10} \\ \nu_{01} \end{bmatrix} \quad (5)$$

The gradients  $\dot{\xi}_{00}(t)$ ,  $\xi_{10}(t)$  and  $\xi_{01}(t)$  in (5) are estimated by finite differences of the acoustic field on the sensor grid comprising four microphones in the configuration illustrated in Figure 1 [14] and computed in a micropower mixed-signal VLSI architecture [16]. The mixing coefficients in (5) represent the corresponding 3-D direction cosines in terms of inter-temporal differences and uniquely determine the direction of the source  $l$ . Therefore, in this representation a system composed by integration of gradient flow and linear instantaneous ICA, can be used to achieve multiple source separation and localization of traveling source signals.

## B. Independent Component Analysis

For this mixing model (1), ICA solution is formulated as a linear transformation that minimizes the statistical dependence between components in the output signals  $\mathbf{y}$

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t), \quad (6)$$

where  $\mathbf{W}$  is  $N \times M$  dimensional unmixing matrix. The unmixing matrix  $\mathbf{W}$  is not uniquely defined, with ambiguity in scaling and permutation. The energy of the source signals cannot be determined, since both  $\mathbf{s}$  and  $\mathbf{A}$  are unknown and any scalar multiplier in one of the sources could be canceled by dividing the corresponding column of  $\mathbf{A}$  by the same scalar. As the order of the independent sources is not predefined, any permutation of the columns of the matrix  $\mathbf{W}$  is a valid solution of the separation problem.

**1) ICA Learning Rules**—Most of the ICA learning algorithms proposed in the literature are based on optimizing a cost function defined as the measure of independence between the components of the output signals [19]. Different approaches, like maximization of entropy [20], minimization of mutual information of the output signals [19], [21] and the maximization of likelihood function [22], lead to the same form of the cost function

$$L(\mathbf{W}) = -N \log(|\det(\mathbf{W})|) - \sum_{i=1}^N \sum_{k=1}^K \log(p_i(y_i(k))), \quad (7)$$

where  $p_i(y_i)$  are the marginal probability density functions (pdfs) of output signals. Term  $\det(\mathbf{W})$  represent the volume conserving property of the linear transformation [23]. ICA learning rule is derived by applying the stochastic gradient descent to (7)

$$\Delta \mathbf{W} = \mu ([\mathbf{W}^T]^{-1} - f(\mathbf{y})\mathbf{x}^T), \quad (8)$$

where  $f(\mathbf{y})$  is the cumulative distribution function

$$f(y_i) = -\frac{dp_i(y_i)/dy_i}{p_i(y_i)}. \quad (9)$$

This update rule was first derived as *InfoMax* learning rule in [20] by maximization of entropy of transformed output signals. Selection of  $f(\mathbf{y})$  as a non-linearity that approximately matches the input cdf's does not affect the performance of the algorithm. The uniform and robust convergence is obtained by using Amari's natural gradient [24], which has the simple form in the space of matrices,  $\mathbf{W}\mathbf{W}^T$ . Multiplying (8) by  $\mathbf{W}\mathbf{W}^T$  leads to the learning rule without matrix inversion

$$\Delta \mathbf{W} = \mu [\mathbf{I} - f(\mathbf{y})\mathbf{y}^T] \mathbf{W}. \quad (10)$$

The convergence of (10) implies  $E\{f_i(y_i)y_i\}=1$  as a constraint on the reconstructed signals. To avoid numerical instability due to non-stationarity in the sources, the Cichocki-Unbehauen (C-U) algorithm [25] introduces a non-holonomic constraint in the natural gradient learning rule (10), by fixing the diagonal terms of the unmixing matrix  $\mathbf{W}$ :

$$\Delta \mathbf{W} = \mu [\Lambda - f(\mathbf{y})\mathbf{y}^T] \mathbf{W}, \quad (11)$$

where  $\Lambda$  is a diagonal scaling matrix. Convergence of the C-U algorithm implies  $\Lambda_{ij} = E[f(y_i)y_j]$ .

The similar algorithm can be derived based on non-linear decorrelation that introduces higher-order statistics into the solution method. This first formulation of ICA, inspired by biomimetic principles, was derived by Herault-Jutten (H-J) [26] and was based on a feedback network topology

$$\mathbf{y} = -\mathbf{W}\mathbf{y} + \mathbf{x}, \quad (12)$$

with zero diagonal terms ( $w_{ii} \equiv 0, \forall i$ ). An independence criterion with nonlinear correlation between output signals yields the on-line learning rule for the off-diagonal terms

$$\Delta w_{ij} = -\mu f(y_i)g(y_j), \quad i \neq j \quad (13)$$

where  $f(\cdot)$  and  $g(\cdot)$  are appropriately chosen, odd-symmetric functions. Good example is the function that matches cumulative distribution function of source signals for  $f$  and linearity  $g(y) \equiv y$  for function  $g$ .

**2) General Outer-Product Learning Rule**—Efficient implementation of ICA in feedforward parallel architecture requires casting learning rules (10) and (13) in a form of the outer-product update rule. H-J learning rule (13) is in the form of the outer-product, however the learning rule is defined for recurrent architecture (12). To map the recurrent architecture onto a feedforward form, we apply the following approximation:

$$\mathbf{y} = (\mathbf{I} + \mathbf{W})^{-1} \mathbf{x} \approx (\mathbf{I} - \mathbf{W}) \mathbf{x} \quad (14)$$

In other words, we choose to implement the H-J rule with linear feedforward network of the type  $\mathbf{y} = \mathbf{W}\mathbf{x}$  with fixed diagonal terms  $w_{ii} \equiv 1$ , and with off-diagonal terms adapting according to (13) [17], [18]. Equivalently, the implemented update rule can be seen as the gradient of *InfoMax* (8) multiplied by  $\mathbf{W}^T$ , rather than the natural gradient multiplication factor  $\mathbf{W}^T \mathbf{W}$ . Interestingly, in the special case of a  $2 \times 2$  network (2 sources and 2 observations) the update rule (13) reduces to non-holonomic (zero-diagonal) form of the C-U rule (11).

To obtain the natural gradient update rule (10) in outer-product form, it is necessary to include a back-propagation path in the network architecture to implement the vector contribution  $\mathbf{z} = \mathbf{W}^T \mathbf{y}$ . The learning rule can then be represented in the form of the decay term and the outer-product term

$$\Delta w_{ij} = \mu w_{ij} - \mu f(y_i) z_j \quad (15)$$

Through quantization of the vector terms in the outer-product rules (13) and (15), as well as the quantization of the decay term in (15), the update rules in the proposed implementation are simplified to the discrete counting operations. In the case of speech signals that are approximately Laplacian distributed, the optimal nonlinear scalar function  $f(y_i)$  can be approximated by  $\text{sign}(y_i)$ , which requires a single bit quantization. Conversely, vectors  $\mathbf{y}$  in rule (13) and  $\mathbf{z}$  in rule (15) are approximated by a 3-level staircase function  $(-1, 0, +1)$  using 2-bit quantization.

### III. Chip Architecture

The functional block diagram of a  $3 \times 3$  outer-product mixed-signal ICA architecture is shown in Figure 2. An analog datapath directly interfaces with input signals and provides analog output signals, without the need for analog-to-digital conversion at the input and digital-to-analog conversion at the output. The digital adaptation offers the flexibility in selection of the learning rules. The mixed-signal architecture is implemented using fully differential switched-capacitor (SC) sampled-data circuits. The coefficients of the unmixing matrix are stored digitally in each cell of the architecture. The update is performed locally by once or repeatedly incrementing, decrementing or holding the current value of counter based on the on-chip defined learning rules or through a range of learning rules served by the external logic. Correlated double sampling performs common mode offset rejection and  $1/f$  noise reduction.

The differential analog input channels directly interface the gradient output signals from a previously developed gradient flow processor for acoustic localization [16], to extend the system functionality to joint separation and localization of up to three acoustic sources. The digital values of the unmixing coefficients correspond to the directional angles of the source signals in the gradient flow representation.

#### A. Vector Matrix Multiplication

ICA is a linear transformation and the main functionality of the proposed implementation is a  $(3 \times 1)$ – $(3 \times 3)$  vector-matrix multiplication. In addition to the vector-matrix multiplication in the computation of the output signals  $\mathbf{y} = \mathbf{W}\mathbf{x}$ , the vector-matrix multiplication is required for the implementation of the natural gradient learning rule in the outer-product form (15),  $\mathbf{z} = \mathbf{W}^T \mathbf{y}$ . Since the same matrix  $\mathbf{W}$  is used in both multiplication operations, the vector  $\mathbf{z}$  is computed using the same circuit used to compute the vector  $\mathbf{y}$ , but with the time-multiplexed input **in** and output **out** signals of the vector-matrix multiplication circuitry. That is, the product  $\mathbf{z} = \mathbf{W}^T \mathbf{y}$  in the natural-gradient learning rule is obtained after computation of the output signal  $\mathbf{y}$ .

The switched-capacitor implementation of the vector-matrix multiplication is shown in Figure 3. The input vector is analog signal, while the coefficients of the unmixing matrix are adaptive digital values. Looking at the product  $\mathbf{y} = \mathbf{W}\mathbf{x}$ , one component of the output signal  $y_i$  is decomposed in differential form as a linear sum of the weighted differential input contributions

$$y_i^+ = - \sum_{j=1}^3 (w_{ij}^+ x_j^- + w_{ij}^- x_j^+)$$

$$y_i^- = - \sum_{j=1}^3 (w_{ij}^+ x_j^+ + w_{ij}^- x_j^-). \quad (16)$$



The multiplication is performed through multiplying D/A capacitor arrays and output of each of these DACs represents product  $w_{ij}^+x_j$  or  $w_{ij}^-x_j$ . The output signal is computed by accumulating outputs from all the DAC cells in the  $i^{th}$  row. The clocks  $\phi_1$  and  $\phi_2$  are non-overlapping, and  $\phi_{1e}$  replicates  $\phi_1$  with its falling edge slightly preceding the falling edge of  $\phi_1$ . All the switches are complementary transmission gate FETs, except the switches controlled by  $\phi_{1e}$ , which are  $n$ -channel FETs. In the precharging phase  $\phi_1$ , all the unit capacitors in the array are precharged to the zero-level reference voltage  $V_{mid}$  (set to  $V_{dd}/2$ ), as well as the feedback capacitor  $C_2$ . The inverters are reset. In the computation phase  $\phi_2$ , the input signals are sampled on unit capacitors and the accumulation is performed on  $C_2$  by high-gain amplifier yielding the valid output signals during  $\phi_2$  phase. The implementation and layout of the multiplying capacitor arrays are the same as in gradient flow acoustic localizer chip [16]. The value of the unit capacitor in the array is 5.1 fF. The value of the feedback capacitor  $C_2$  is externally set to accommodate variable gain in the vector-matrix multiplication and can take one of the four values: 0.5 pF, 1 pF, 2 pF and 4 pF. The vector-matrix multiplier is followed by the sample-and-hold circuit that provides the analog output signal.

A cascoded pseudo-nMOS inverter biased in weak inversion (the upper range of the subthreshold regime) is used as high-gain amplifier [16]. The choice of telescopic operational amplifier without tail transistor is driven by smaller area and reduced noise and power dissipation [27], [28], while the weak inversion is chosen as the region of operation due to extended output dynamic range and the maximum transconductance-to-current ratio for maximum energy efficiency at the highest possible speed [29]. The performance of the designed cascoded amplifier was simulated at 200 nA of biasing current with the load capacitance of 1 pF. At 3 V supply, simulations indicate an open-loop dc gain of 91 dB and gain-bandwidth product of 844.3 kHz. The biasing current in each of the amplifiers in the proposed architecture was set by considerations of sampling frequency, slew-rate, and power dissipation.

## B. Comparator Design

The discrete values of functions  $\text{sgn}(y_i)$ ,  $\text{quant}(y_j)$  and  $\text{quant}(z_j)$  in (15) are obtained through level comparisons. Implementation of the comparator, that is able to compare signal with variable level, consists of the preamplifier, shown in Figure 4, followed by a latched, regenerative comparator [16]. The clocks  $\phi_1^c$  and  $\phi_2^c$  are non-overlapping and their relative timing with respect to clocks  $\phi_1$  and  $\phi_2$  is shown as inset in Figure 4, as well as the changes in threshold voltage  $V_{TH}$  for 3-level comparison.  $\phi_{1e}^c$  replicates  $\phi_1^c$  with its falling edge slightly preceding the falling edge of  $\phi_1^c$ . The value of capacitors  $C_1$  and  $C_c$  is 300 fF.

While the output signals are valid,  $y_i^{\pm}$  is sampled in phase  $\phi_1^c$  on capacitor  $C_1$ . The sign of the comparison of  $y_i$  with variable level threshold  $V_{TH}$  is computed in the evaluate phase  $\phi_2^c$ , through capacitive coupling into the amplifier input node. The change in voltage  $V_{TH}$  in phase  $\phi_2^c$  will lead to multiple level comparisons in a single clock cycle. The offset of the comparator was measured to be 10 mV and is consisted for the different level comparisons.

### C. Learning Rule

The parallel architecture with the representation of the ICA learning rules in the form of outer-product enables the local updates of coefficients of the unmixing matrix  $\mathbf{W}$ . The coefficients are represented as 14-bit values in two's complement and are stored in a counter. In the gradient flow framework, the unmixing coefficients yield the 3-D direction cosines in terms of inter-temporal differences  $\tau_1^\ell$  and  $\tau_2^\ell$ . The update is performed by incrementing, decrementing or holding the current value of the counter.

The implementation of H-J learning rule (13) is performed as [17]

$$w_{ij}^+[n+1]=w_{ij}^+[n] - \text{sgn}(y_i[n])\text{quant}(y_j[n])$$

$$w_{ij}^-[n+1]=2^{14} - 1 - w_{ij}^+[n+1], \quad (17)$$

where  $\text{quant}(y_j)$  is a 3-level staircase function  $(-1,0,+1)$ , that is coded using 2-bits,  $\text{sgn}$  and  $\text{mag}$ , as illustrated in Figure 5. The natural gradient learning update (15) comprises the outer product update

$$w_{ij}^+[n+1]=w_{ij}^+[n] - \text{sgn}(y_i[n])\text{quant}(z_j[n])$$

$$w_{ij}^-[n+1]=2^{14} - 1 - w_{ij}^+[n+1], \quad (18)$$

and periodical update that is proportional to the value of the first two MSBs of  $w_{ij}$ . Additional levels could be introduced in a staircase function  $\text{quant}(y_j)$  by increasing a number of levels in the comparator signal  $V_{TH}$ , as described in the Section III-B, at the penalty of increased power consumption due to additional clock transitions. The 8 most significant bits of the weights are presented to multiplying D/A capacitor array in thermometer code to construct the output signals. The remaining 6 bits in the coefficient registers provide flexibility in programming the update rate to tailor convergence.

## IV. Experimental Results

A prototype  $3 \times 3$  mixed-signal ICA processor was designed, fabricated, and tested. The architecture is integrated on a single  $3\text{mm} \times 3\text{mm}$  chip fabricated in  $0.5 \mu\text{m}$  3M2P CMOS technology. Figure 6 shows the micrograph of the chip.

All the experiments with the chip were conducted at 3 V supply voltage and with the zero-level reference  $V_{mid}$  set to 1.5 V. The signal swing at the chip inputs is 2.4 V peak-to-peak. The reference voltage  $V_{mid}$  and three biasing voltages for cascoded inverter were generated on-board for testing purposes. However, they might have been generated internally at expense of small increase in power consumption.

The integral nonlinearity(INL) of the multiplying DAC is measured at 0.54 LSB. The effective resolution in the output signal as the result of the vector-matrix multiplication is 8.83 bits. The measured power consumption of the chip is 195  $\mu$ W at 16 kHz sampling rate.

The digital estimates of coefficients  $w_{ij}$  in unmixing matrix are obtained directly from the counters at convergence and are output 4-bits at time. The chip also outputs the estimated source signals  $y_i(t)$ . The output signals are presented in complementary analog format through sample-and-hold buffers.

### A. Synthetic Speech Experiments

For the full characterization of the separation performance of ICA processor, the synthetic mixtures of the speech signals were generated under diverse conditions in a controlled environment. The observed signals were artificially generated as received on the four microphone array in a configuration illustrated in Figure 1. The distance between microphones was set at 1 cm with sampling frequency of 16 kHz. The source signals were speech signals from TIMIT database. The microphone signals were first processed by acoustic localizer [16] that computes the spatial gradient signals (5) using on-grid approximation. The spatial gradients are the input signals to the ICA processor. ICA processor was configured to implement the outer-product update algorithms in (13) and (15). The quantization level in 3-level approximation of  $\text{quant}(\mathbf{y})$  was set to 180 mV amplitude change in voltage  $V_{TH}$ , while the level change in  $\text{quant}(\mathbf{z})$  was 150 mV throughout all the experiments.

In the on-chip implementation of the ICA learning rules, we introduced quantization of the vector terms in the learning rules (13) and (15), as well as quantization of the decay term in (15). To assess the effect of the quantization on the separation performance, ICA algorithms implementing H-J learning rule (13) and natural gradient learning rule (15) without quantization were implemented in MATLAB. As additional benchmark for the separation performance, the efficient FastICA(EFICA) algorithm [30] was used. As a measure of separation performance, we used the signal-to-interference ratio (SIR). SIR in a single output  $y_i$  is computed as

$$\text{SIR}_i = -10 \log_{10} \frac{\sum_j \langle y_{ij}^2 \rangle - \max_j \langle y_{ij}^2 \rangle}{\max_j \langle y_{ij}^2 \rangle}, \quad (19)$$

where  $y_{ij}$  is the contribution of the signal  $j$  to the output signal  $i$ . The reported SIR for the separation of multiple sources is obtained as the minimum of SIR in all output signals

$$\text{SIR} = \min_i \text{SIR}_i. \quad (20)$$

In the first experiment, the effect of the angular distance between two sources on the separation performance was examined. The same elevation angles were assumed for both sources and the azimuth angle of the first source was set at  $\theta_1=30^\circ$ . The azimuth angle of the

second source  $\theta_2$  was varied from  $-15^\circ$  to  $135^\circ$  in increments of  $15^\circ$ . We have omitted the locations of the second source at  $0^\circ$  and  $90^\circ$  where the separation is trivial. The separation performance of ICA online on-chip rules (13) (on-chip HJ) and (15) (on-chip NG), as well as the performance obtained from MATLAB with H-J, natural gradient (NG) and EFICA algorithm are shown in Figure 7. The results demonstrate a uniform separation performance for EFICA, while performance of on-line learning rules slightly depends on the directional separation. HJ online learning rule has demonstrated a lower convergence rate than natural-gradient on-line rule, with also higher sensitivity to the initial conditions. The experiment was repeated with three sources present in the environment, with incidence angle corresponding to  $30^\circ$ ,  $70^\circ$  and  $135^\circ$ . The simulated SIRs in three channels for on-chip NG were 28.5 dB, 25.6 dB and 18.2 dB. We also assessed the effect of the finite resolution of the vector-matrix multiplication on the separation performance of the quantized on-chip learning rules through MATLAB simulations. Increasing the effective resolution of the vector-matrix multiplication beyond 9 bits did not result in any significant improvement in the separation performance.

In the second experiment, the effect of the acquisition noise on the separation performance is investigated. We assumed that a white, spatially uncorrelated Gaussian noise sources are added to each sensor. The results for different signal-to-noise ratios (SNR) are presented in Figure 8. In the noisy environment, on-line quantized learning rules outperform EFICA.

In the mixing model adopted in the gradient flow representation (3), the anechoic environment is assumed. However, in the real room environment, due to reverberations the observed microphone signal is a sum of multi-path replicas of the source signal, that is a sum of time-delayed and attenuated source signals, where the delays and attenuations depend on the room geometry and the reflection coefficient of the walls. To measure the influence of the reverberations on the separation performance, we model a real room environment using image model [31]. The room dimensions were selected as [6m, 4m, 3m] and the reflection coefficient was the same for all the room surfaces. The location of the speakers and the microphone array is illustrated in Figure 9. The reflection coefficient was varied to determine the effect of the reverberation on the separation performance. The separation, as the function of the reflection coefficient, is shown in Figure 10. The separation performance degrades with the increase of the reflection coefficient, but satisfying performance is demonstrated in mild reverberant conditions. The separation is sustained as long as the direct path signal is stronger than the multi-path signals. The reflection coefficient of 0.52 in the presented room corresponds to the reverberation time of 300 ms.

## B. Room Speech Separation Experiments

The separation performance of ICA processor was demonstrated and characterized in a real room environment. The recordings were performed in a typical conference room with size that corresponds to the simulated room in the synthetic room experiments. The input signals to the ICA processor were the spatial gradient signals, that are outputs of the gradient flow acoustic localizer [16]. The microphone array comprises four omnidirectional miniature microphones (Knowles FG-3629) in configuration shown in Figure 1. Microphones were arranged in a circular array with radius of 0.5 cm. The sensitivity of microphones is  $-53$  dB

and typical noise level is at 27 dB. The microphone signals were passed through a second-order bandpass filter with low-frequency cutoff set at 130 Hz and high-frequency cutoff set at 4.3 kHz. The signals were also amplified by a factor of 20. The system sampling frequency was set to 16 kHz. The speech signals, same as the ones used in the synthetic experiments, were presented through loudspeakers positioned at 1.5 m distance from the array.

To provide the quantifiable separation performance, speech segments were presented individually through either loudspeaker at different time instances. The first source was kept at the azimuth angle of  $30^\circ$ , while the second was moved from  $45^\circ$  to  $165^\circ$  in increments of  $15^\circ$ . The two recorded datasets were then added, and presented to the gradient flow localizer ASIC [16]. The gradient signals obtained from the chip were then presented to the ICA processor, configured to implement the outer-product update algorithms in (13) and (15). The obtained SIRs for both on-chip learning rules, as well as for software H-J, natural gradient and EFICA ICA algorithms, are presented in Figure 11. The real room SIRs are in accordance with the simulation results and verify the separation performance in a mild reverberant environment. The experiments were repeated with joint presentation of both sources and 14-bit digital weights that represent the coefficients of the unmixing matrix were recorded. The angles of incidence of the sources relative to the array were derived and compared to the localization results obtained by acoustic localizer ASIC when a single source was presented. The angles obtained through LMS bearing estimation under individual source presentation are within  $2^\circ$  to the angles produced by both on-line ICA learning rules. In Figure 12, the input and the output signals of the ICA processor, that is the spatial gradients (3) and the reconstructed source signals, are presented for azimuth angles of  $\theta_1=30^\circ$  and  $\theta_2=105^\circ$ . We verified that the time-waveforms of the residue signals in each of the reconstructed outputs,  $y_{12}$  and  $y_{21}$ , are free of direct path signal components and are dominated by multipath contributions due to room reverberation. The performance of the proposed ICA mixed-signal implementation has been compared to the implementations in the digital-domain in Table I. It is important to note that the proposed IC has analog input and output signals which eliminates the need for data conversion. This results in significantly lower power consumption and form factor of the corresponding sensory system for source separation compared to the system that contains digital ASIC. Additionally, the operational frequency in the proposed IC matches the bandwidth of the input signals, which is not the case for digital ASICs.

## V. Conclusions

The proposed parallel low-power ICA architecture with outer-product learning rules using the gradient flow representation demonstrated state-of-the-art separation performance in noisy and mild reverberant conditions. The architecture mitigates the need for input or output data conversion as the inputs are the analog observed signals and the outputs are the analog reconstructed source signals with digitally stored coefficients of the unmixing matrix. The measured characteristics of the fabricated chip are summarized in Table II. These results suggest application of the gradient flow system integrating spatial sensing and static ICA with miniature microphone arrays to intelligent hearing aids with adaptive suppression of interfering signals and nonstationary noise. The separation performance of the proposed

system can be extended to moderate reverberation environment by using subband decomposition of spatial gradient signals and static ICA applied in each frequency band.

## Acknowledgments

This work was supported by the National Science Foundation CAREER Award 0846265. The chip was fabricated through the MOSIS foundry service.

## References

1. Pedersen MS, Larsen J, Kjems U, Para LC. A Survey of Convolutional Blind Source Separation Methods. Springer Multichannel Speech Processing Handbook. 2007:1065–1084.
2. Ferreol A, Albera L, Chevalier P. Fourth order blind identification of underdetermined mixtures of sources (FOBIUM). IEEE Trans. on Signal Processing. 2005; 53(5):1640–1653.
3. Kachenoura A, Albera L, Senhadji L, Comon P. ICA: A potential tool for BCI systems. IEEE Signal Process. Mag. 2008; 25(1):57–68.
4. Cichocki, A., Amari, S. Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. New York: John Wiley; 2002.
5. Cohen MH, Andreou AG. Analog CMOS Integration and Experimentation with an Autoadaptive Independent Component Analyzer. IEEE Trans. Circuits and Systems II. 1995; 42(2):65–77.
6. Gharbi, ABA., Salam, FMA. Implementation and Test Results of a Chip for the Separation of Mixed Signals; Proc. Int. Symp. Circuits and Systems (ISCAS'95); 1995.
7. Kim CM, Park HM, Kim T, Choi YK, Lee SY. FPGA implementation of ICA algorithm for blind signal separation and adaptive noise canceling. IEEE Trans. on Neural Network. 2003; 14(5):1038–1046.
8. Shyu KK, Lee MH, Wu YT, Lee PL. Implementation of pipelined fastICA on FPGA for real-time blind source separation. IEEE Trans. on Neural Network. 2008; 19(6):958–970.
9. Van L-D, Wu D-Y, Chen C-S. Energy-Efficient FastICA Implementation for Biomedical Signal Separation. IEEE Trans. on Neural Network. 2011; 22(11):1809–1822.
10. Yang C-H, Shih Y-H, Chiueh H. An 81.6  $\mu$ W FastICA Processor for Epileptic Seizure Detection. IEEE Trans. on Biomedical Circuits and Systems. 2015; 9(1):60–71. [PubMed: 24968296]
11. Miles RN, Su Q, Cui W, Shetye M, Degertekin FL, Bicen B, Garcia C, Jones S, Hall N. A low-noise differential microphone inspired by the ears of the parasitoid fly *Ormia ochracea*. J. Acoust. Soc. Am. 2009; 125(5):2013–2026. [PubMed: 19354377]
12. Ando, S., Kurihara, T., Watanabe, K., Yamanishi, Y., Ooasa, T. Novel theoretical design and fabrication test of biomimicry directional microphone; International Solid-State Sensors, Actuators and Microsystems Conference TRANSDUCERS 2009; 2009. p. 1932-1935.
13. Hamacher V, Chalupper J, Eggers J, Fisher E, Kornagel U, Puder H, Rass U. Signal Processing in High-End Hearing Aids: State of the Art, Challenges, and Future Trends. EURASIP Journal on Applied Signal Processing. 2005; 18:2915–2929.
14. Cauwenberghs, G., Stana evi , M., Zweig, G. Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'2001). Australia: Sydney; 2001 May 6–9. Blind Broadband Source Localization and Separation in Miniature Sensor Arrays.
15. Stana evi M, Cauwenberghs G, Zweig G. Gradient Flow Adaptive Beamforming and Signal Separation in a Miniature Microphone Array. Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing (ICASSP'2002). 2002; 4:4016–4019.
16. Stana evi M, Cauwenberghs G. Micropower Gradient Flow VLSI Acoustic Localizer. IEEE Transactions on Circuits and Systems I : Regular Papers. 2005; 52(10):2148–2157.
17. Celik, A., Stana evi , M., Cauwenberghs, G. Mixed-signal real-time adaptive blind source separation; Proc. IEEE Int. Symp. Circuits Syst; 2004. p. 760-763.
18. Celik, A., Stana evi , M., Cauwenberghs, G. Adv. Neural Information Processing Systems (NIPS'2005). Vol. 18. Cambridge: MIT Press; 2006. Gradient Flow Independent Component Analysis in Micropower VLSI.



19. Comon P. Independent component analysis, a new concept? *Signal Processing*. 1994; 36:287–314.
20. Bell A, Sejnowski T. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*. 1995; 7:1129–1159. [PubMed: 7584893]
21. Amari, S., Cichocki, A., Yang, H. *Adv. Neural Information Processing Systems*. Vol. 8. Cambridge MA: MIT Press; 1996. A new learning algorithm for blind signal Separation; p. 757-763.
22. Nadal J, Parga N. Non linear neurons in the low noise limit : a factorial code maximizes information transfer. *Network*. 1994; 5:565–581.
23. Obradovic D, Deco G. Information Maximization and Independent Component Analysis: Is There a Difference? *Neural Computation*. 1998; 10(8):2085–2101. [PubMed: 9804672]
24. Cichocki, A., Unbehauen, R., Moszczanski, L., Rummert, E. *Int. Symp. Artificial Neural Networks ISANN-94*. Taiwan: 1994 Dec. A New On-Line Adaptive Learning Algorithm for Blind Separation of Sources; p. 406-411.
25. Cichocki A, Unbehauen R, Rummert E. Robust Learning Algorithm for Blind Separation of Signals. *Electronics Letters*. 30(17):1386–1387.
26. Jutten C, Herault J. Blind Separation of Sources, part I: An Adaptive Algorithm Based on Neuromimetic Architecture. *Signal Proc*. 1991; 24(1):1–10.
27. Nicollini G, Moretti F, Conti M. High-Frequency Fully Differential Filter Usign Operational Amplifier Without Common-Mode Feedback. *IEEE Journal of Solid-State Circuits*. 1989; 24(3): 803–813.
28. Li J, Moon U. A 1.8-V 67-mW 10-bit 100-MS/s pipelined ADC using time-shifted CDS technique. *IEEE Journal of Solid-State Circuits*. 2004 Sep; 39(9):1468–1476.
29. Vittoz, E. *Micropower Techniques*. In: Franca, Tsividis, editors. *Design of Analog-Digital VLSI Circuits for Telecommunications and Signal Processing*. 2nd. Englewood Cliffs, NJ: Prentice-Hall; 1994. p. 53-96.
30. Koldovsky Z, Tichavsky P, Oja E. Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramer-Rao Lower Bound. *IEEE Trans. on Neural Networks*. 2006; 17(5):1265–1277. [PubMed: 17001986]
31. Allen JB, Berkley DA. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Amer*. 1979 Apr.65:943–950.

## Biographies



**Milutin Stana evi** (S'00-M'05) received the B.S. degree in Electrical Engineering from the University of Belgrade, Serbia in 1999. He received the M.S. and Ph.D. degrees in Electrical and Computer Engineering from Johns Hopkins University, Baltimore, MD, in 2001 and 2005, respectively.

In 2005, he joined the faculty of the Department of Electrical and Computer Engineering at Stony Brook University, Stony Brook, NY, where he is currently an Associate Professor. His research interests include mixed-signal VLSI circuits, systems, and algorithms for parallel multi-channel sensory information processing with emphasis on acoustic source separation and breath analysis, micropower biomedical instrumentation and readout ICs for radiation detection.

Dr. Stana evi is a recipient of the National Science Foundation CAREER award and IEEE Region 1 Technological Innovation Award. He is an Associate Editor of the IEEE Transactions on Biomedical Circuits and Systems and serves on several technical committees of the IEEE Circuits and Systems Society.



**Shuo Li** (S'07) received the B.Eng degree from Zhejiang University, Hangzhou, China in 2006, M.Phil degree from The Hong Kong University of Science and Technology, Hong Kong in 2008 and Ph.D. degree from Stony Brook University, Stony Brook, NY in 2015, all in electrical engineering. His research is on acoustic localization and separation and VLSI microsystem design.

Currently, he is an electrical engineer at Second Sight Medical Products, Inc. He works on the research and development of retinal and cortical stimulation medical devices.



**Gert Cauwenberghs** (S'89-M'94-SM'04-F'11) received the Ph.D. degree in electrical engineering from California Institute of Technology, Pasadena, in 1994.

He is Professor of Bioengineering and Co-Director of the Institute for Neural Computation at the University of California-San Diego, La Jolla, CA, USA. He was previously Professor of Electrical and Computer Engineering at Johns Hopkins University, Baltimore, MD, USA, and Visiting Professor of Brain and Cognitive Science at Massachusetts Institute of Technology, Cambridge, MA, USA. He co-founded Cognionics Inc., and chairs its Scientific Advisory Board. His research focuses on micropower biomedical instrumentation, neuronsilicon and brainmachine interfaces, neuromorphic engineering, and adaptive intelligent systems.

Dr. Cauwenberghs received the NSF Career Award in 1997, ONR Young Investigator Award in 1999, and Presidential Early Career Award for Scientists and Engineers in 2000. He serves IEEE in a variety of roles including as General Chair of the IEEE Biomedical Circuits and Systems Conference (BioCAS 2011, San Diego, CA, USA), as Program Chair of the IEEE Engineering in Medicine and Biology Conference (EMBC 2012, San Diego, CA, USA), and as Editor-in-Chief of the Transactions on Biomedical Circuits and Systems.

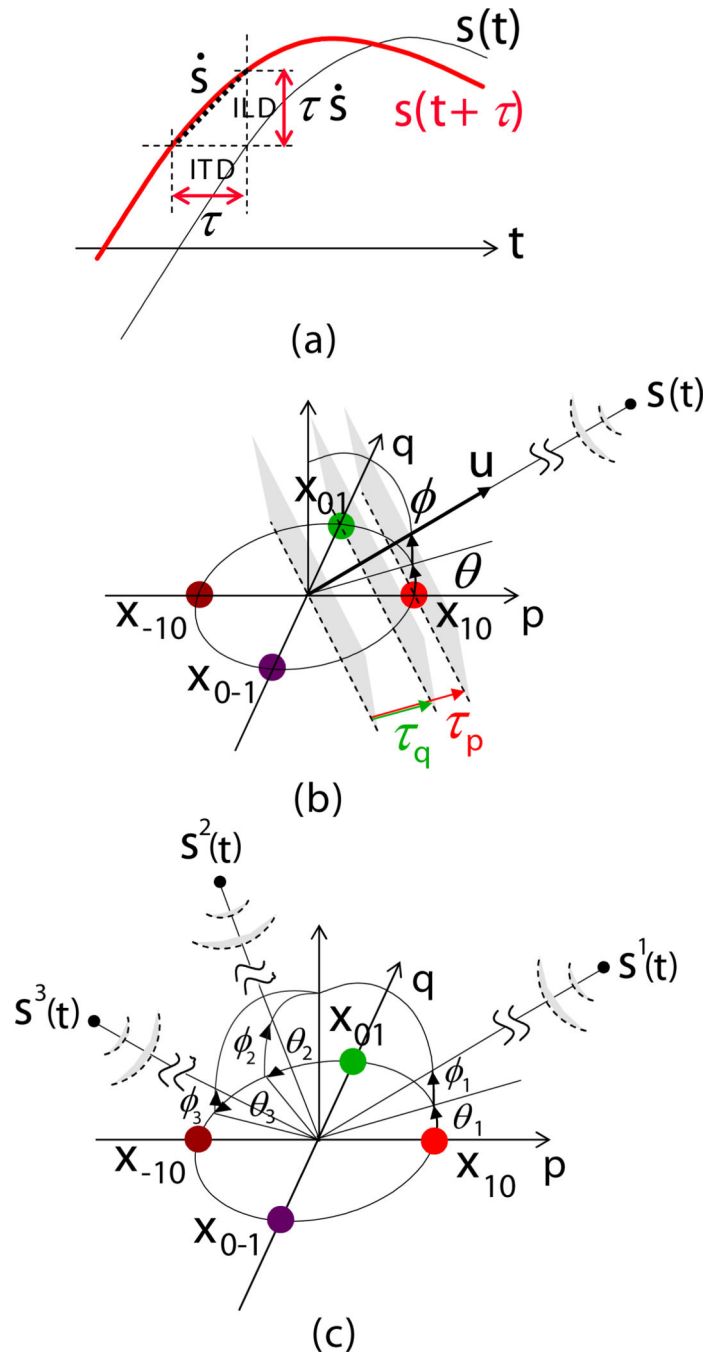


Author Manuscript

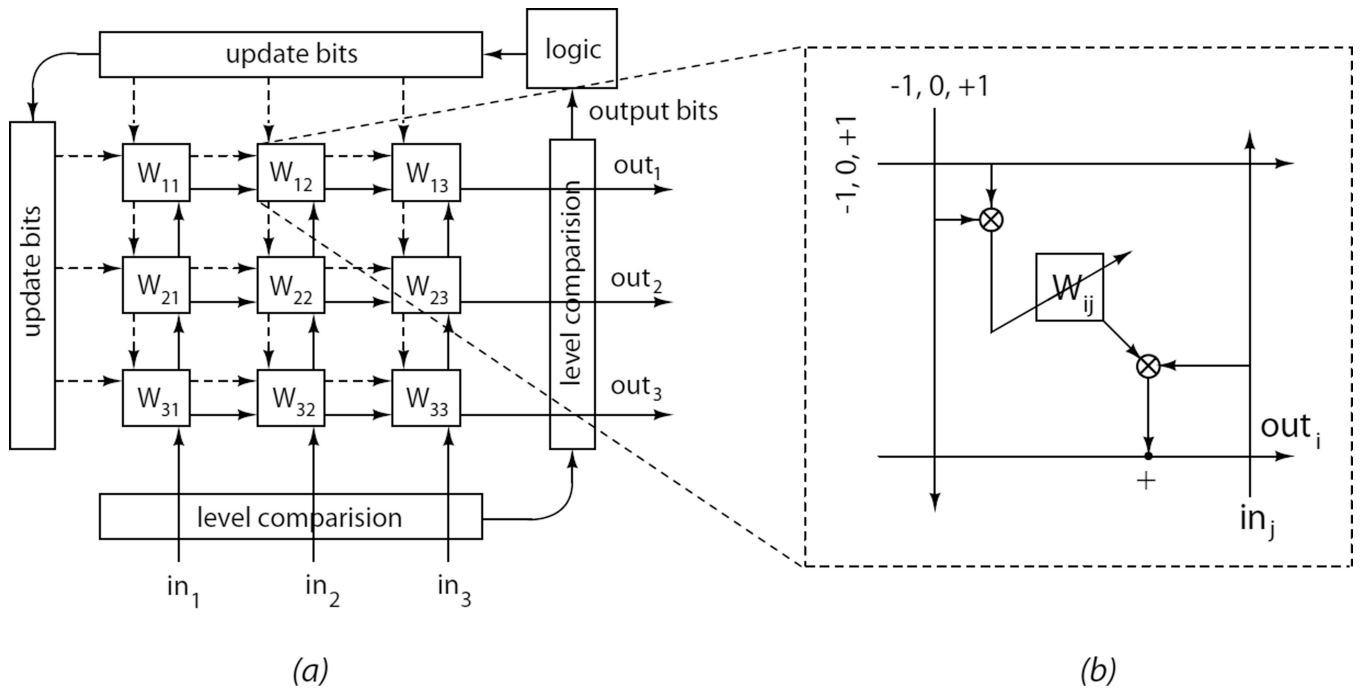
Author Manuscript

Author Manuscript

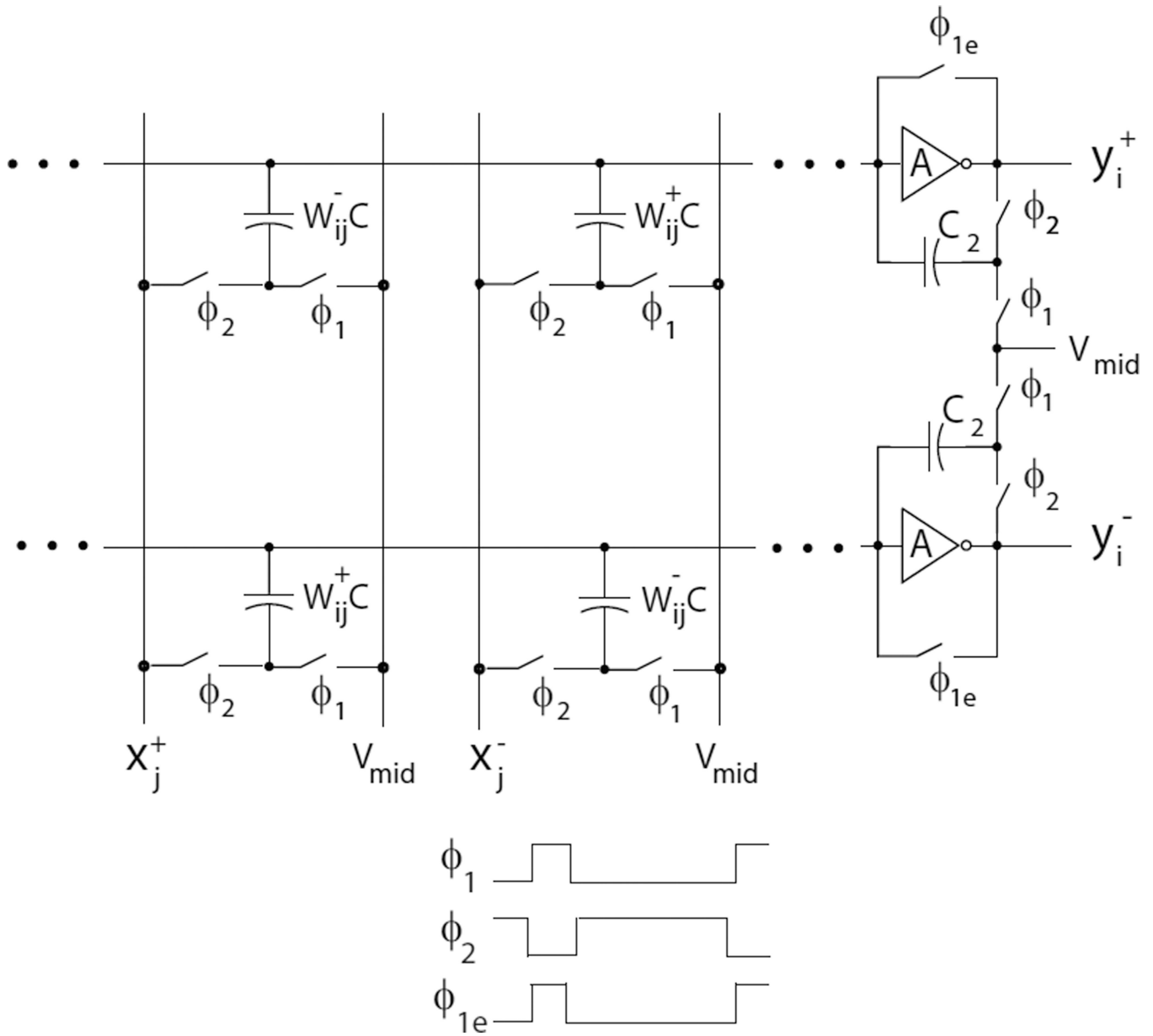
Author Manuscript



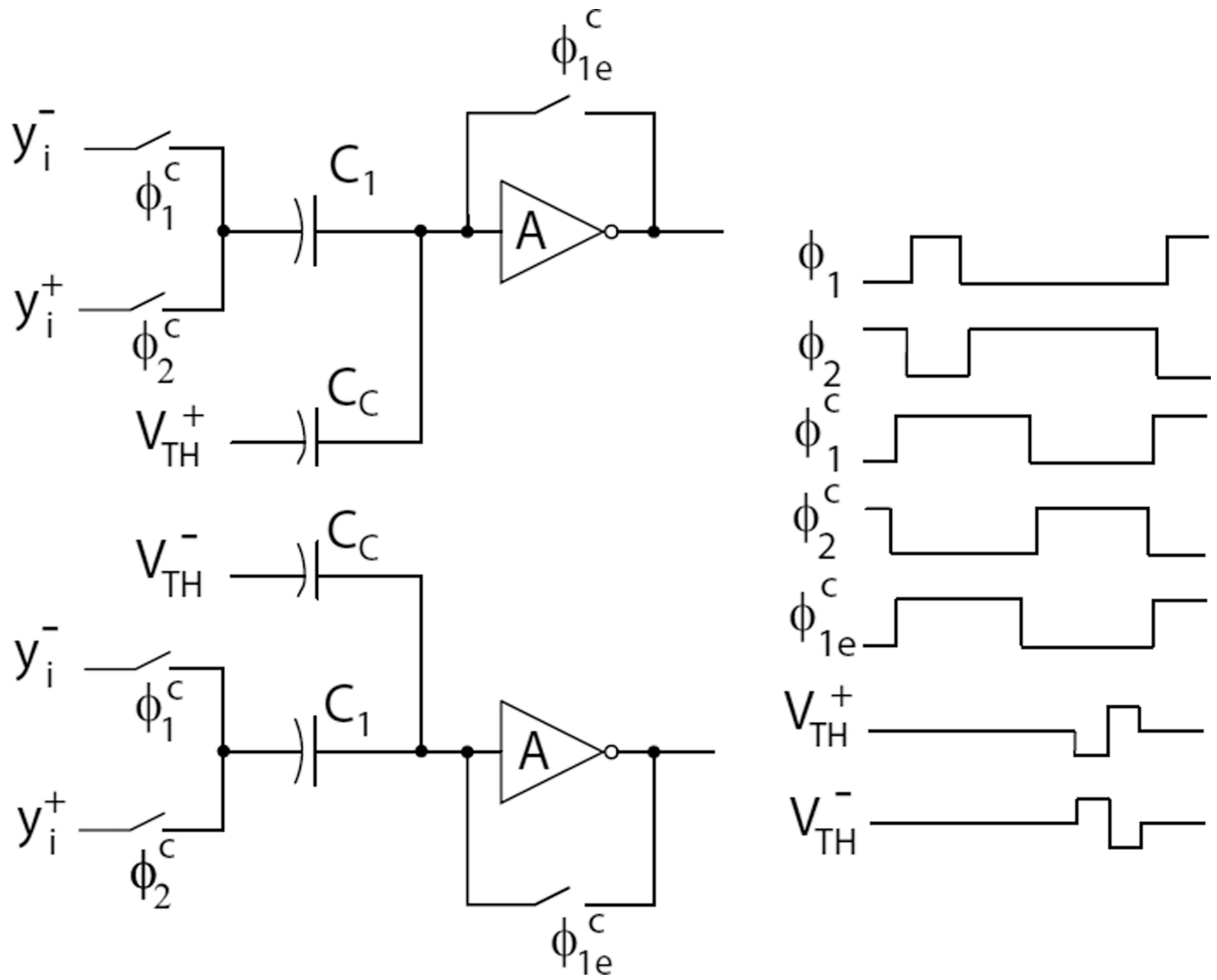
**Fig. 1.** (a) Gradient flow principle. At low aperture, interaural level differences (ILD) and interaural time differences (ITD) are directly related, scaled by the temporal derivative of the signal [18]. (b) Propagation delays  $\tau_p$  and  $\tau_q$  scale the spatial gradients in perpendicular directions. (c) Joint separation and localization of up to three sources  $s_1, s_2$  and  $s_3$  impinging on the array.



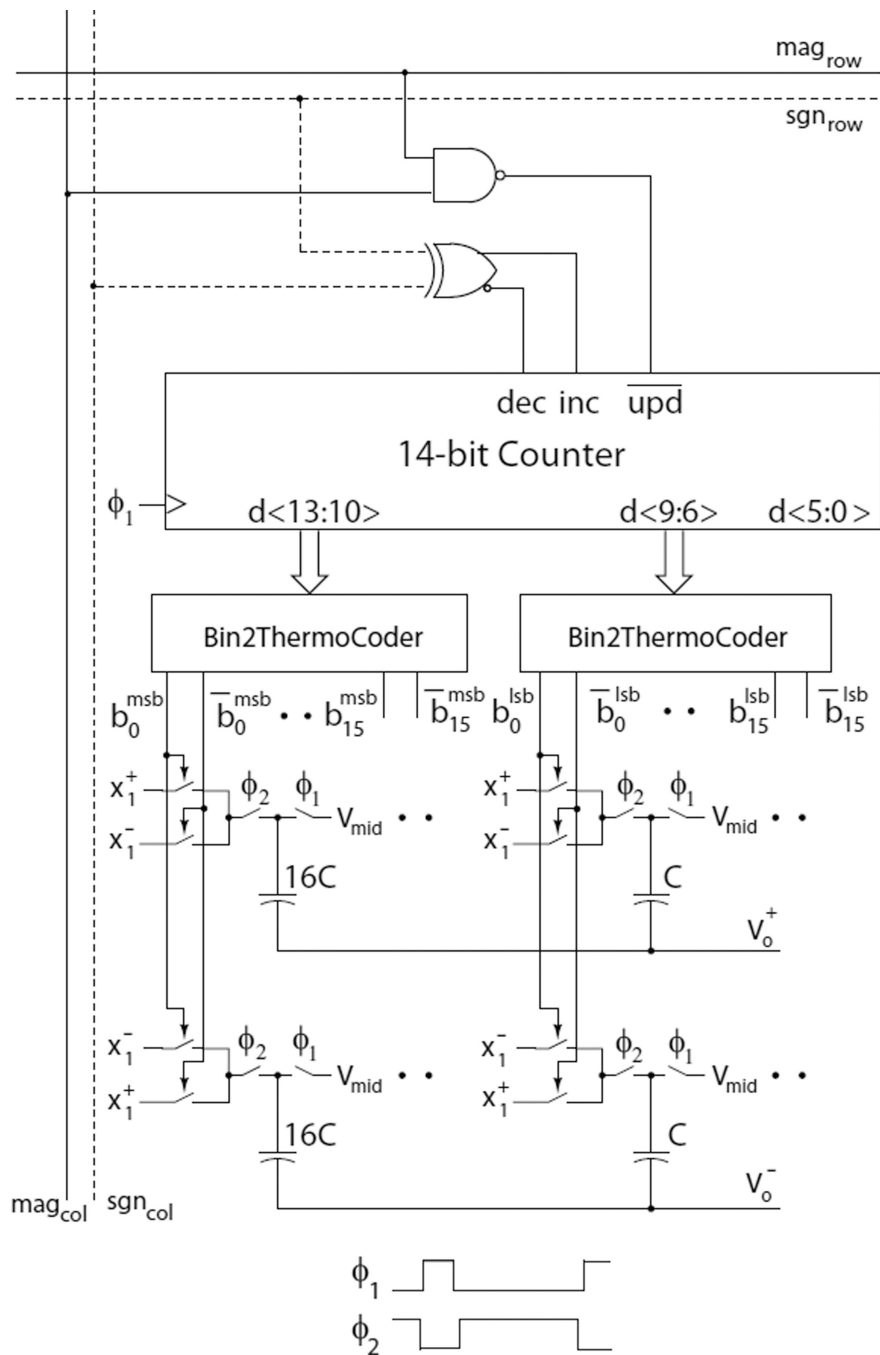
**Fig. 2.** (a) Reconfigurable parallel mixed-signal ICA architecture implementing general outer-product form of ICA update rules with one unmixing coefficient cell shown in (b).



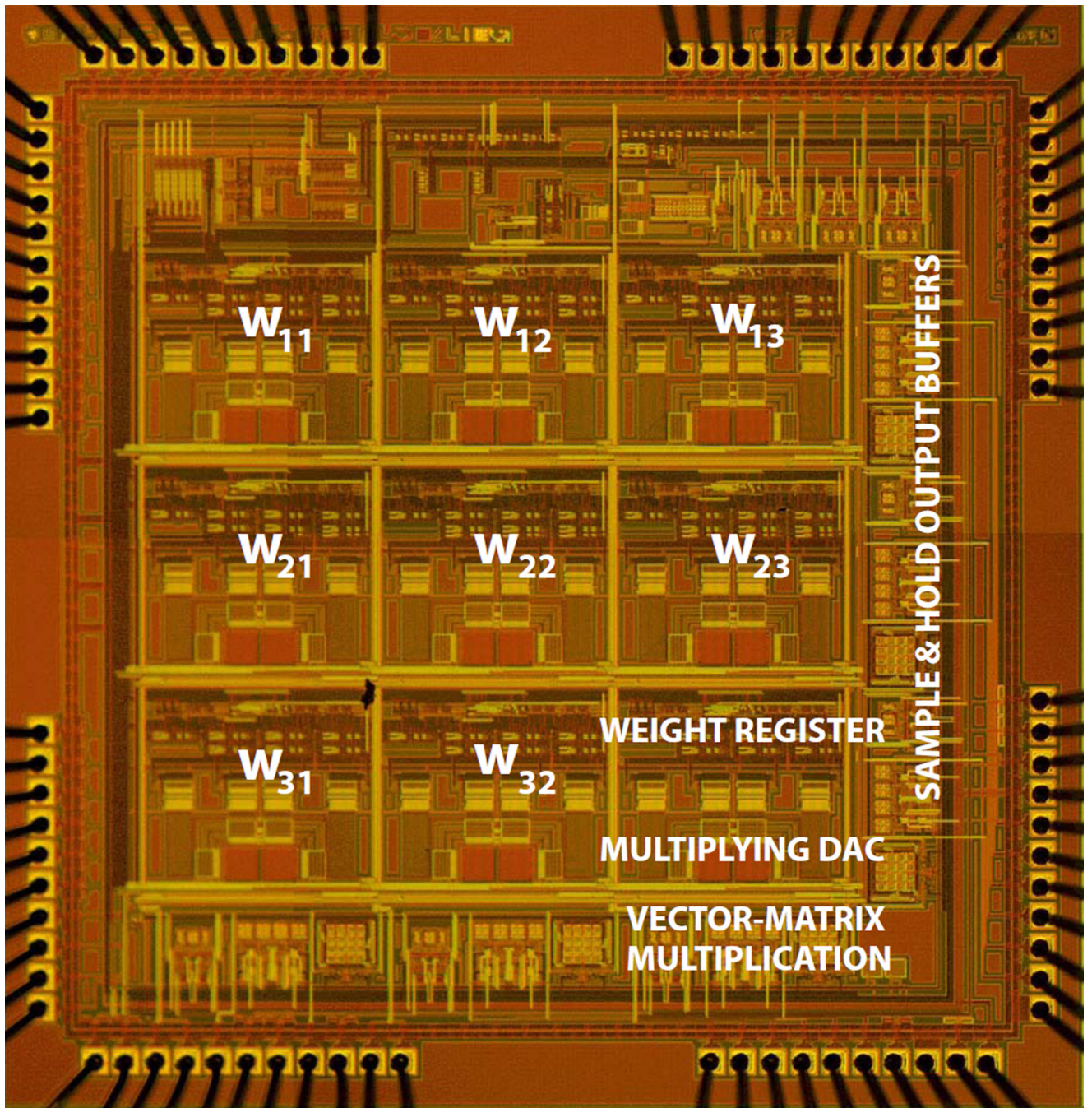
**Fig. 3.** Switched-capacitor implementation of the vector-matrix multiplication for computation of the  $i^{th}$  component of the output signal  $\mathbf{y}$ .



**Fig. 4.** Design of the comparator for sequential comparison of the signals  $y$  and  $z$  with variable voltage levels defined by the threshold voltage  $V_{TH}$ .

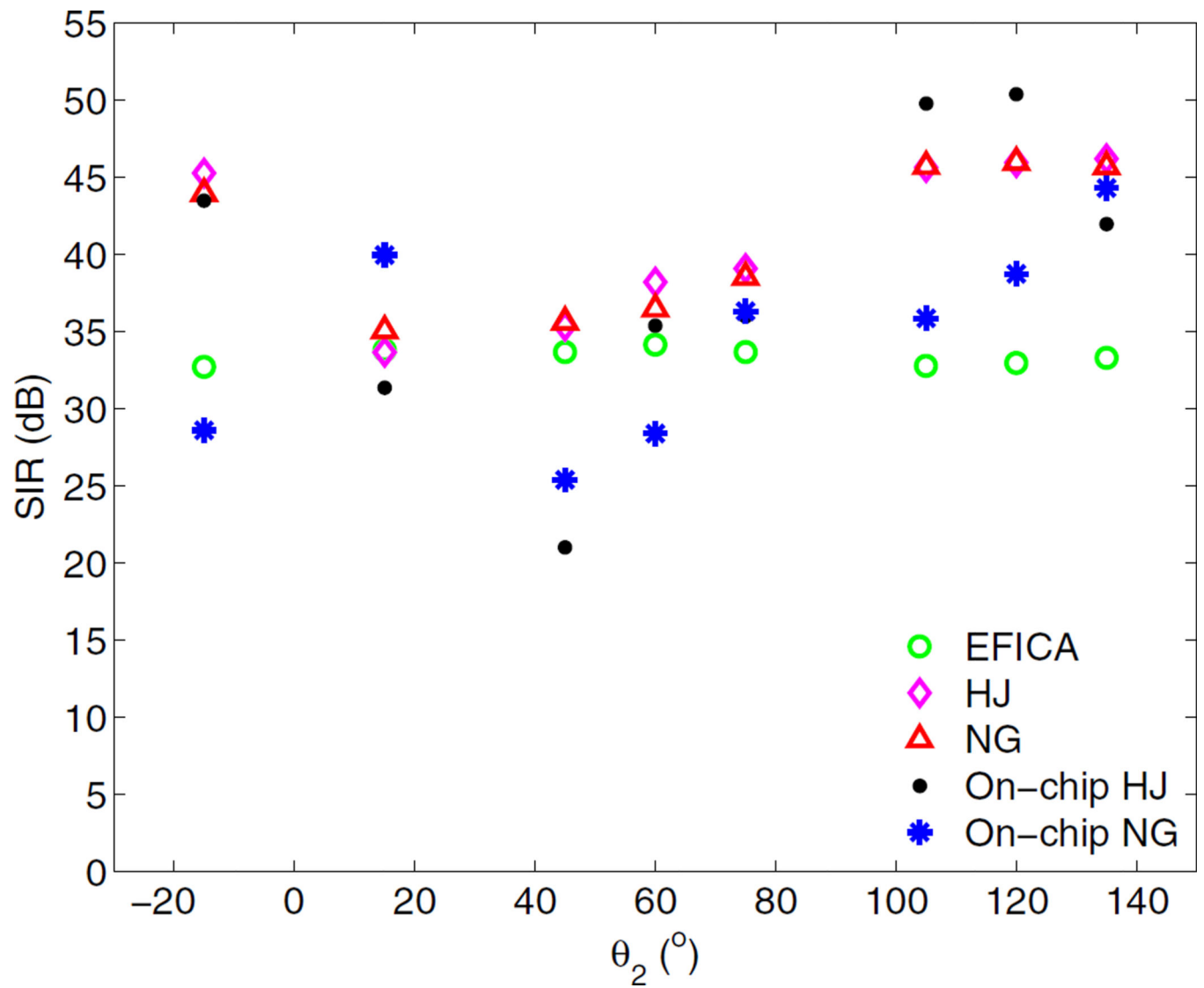


**Fig. 5.** The unmixing coefficient  $w_{ij}$  is stored in a 14-bit counter and presented for vector-matrix multiplication through 8-bit multiplying D/A capacitor array [16].

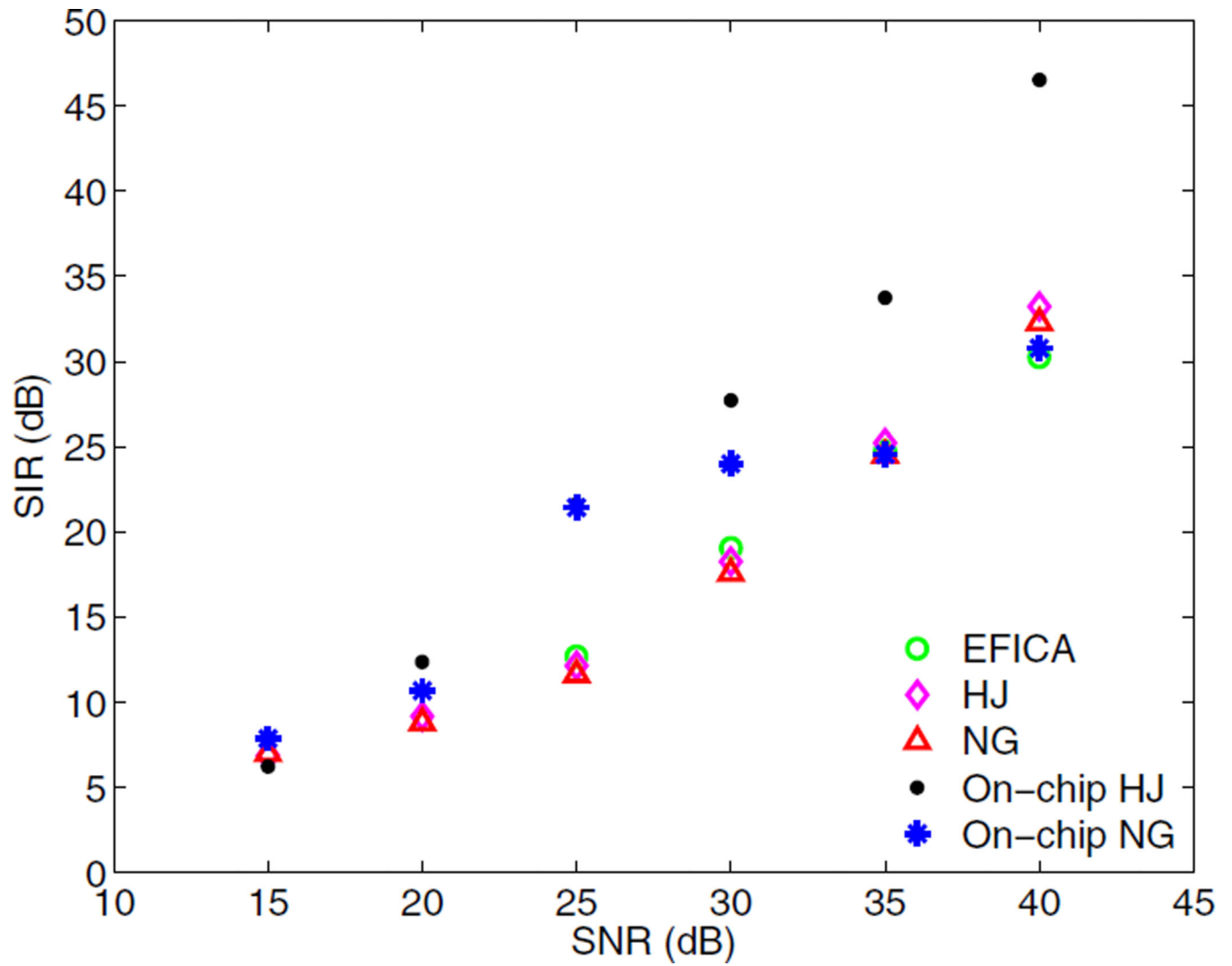


**Fig. 6.**  
Micrograph of 3 mm × 3 mm chip in 0.5 μm CMOS technology.



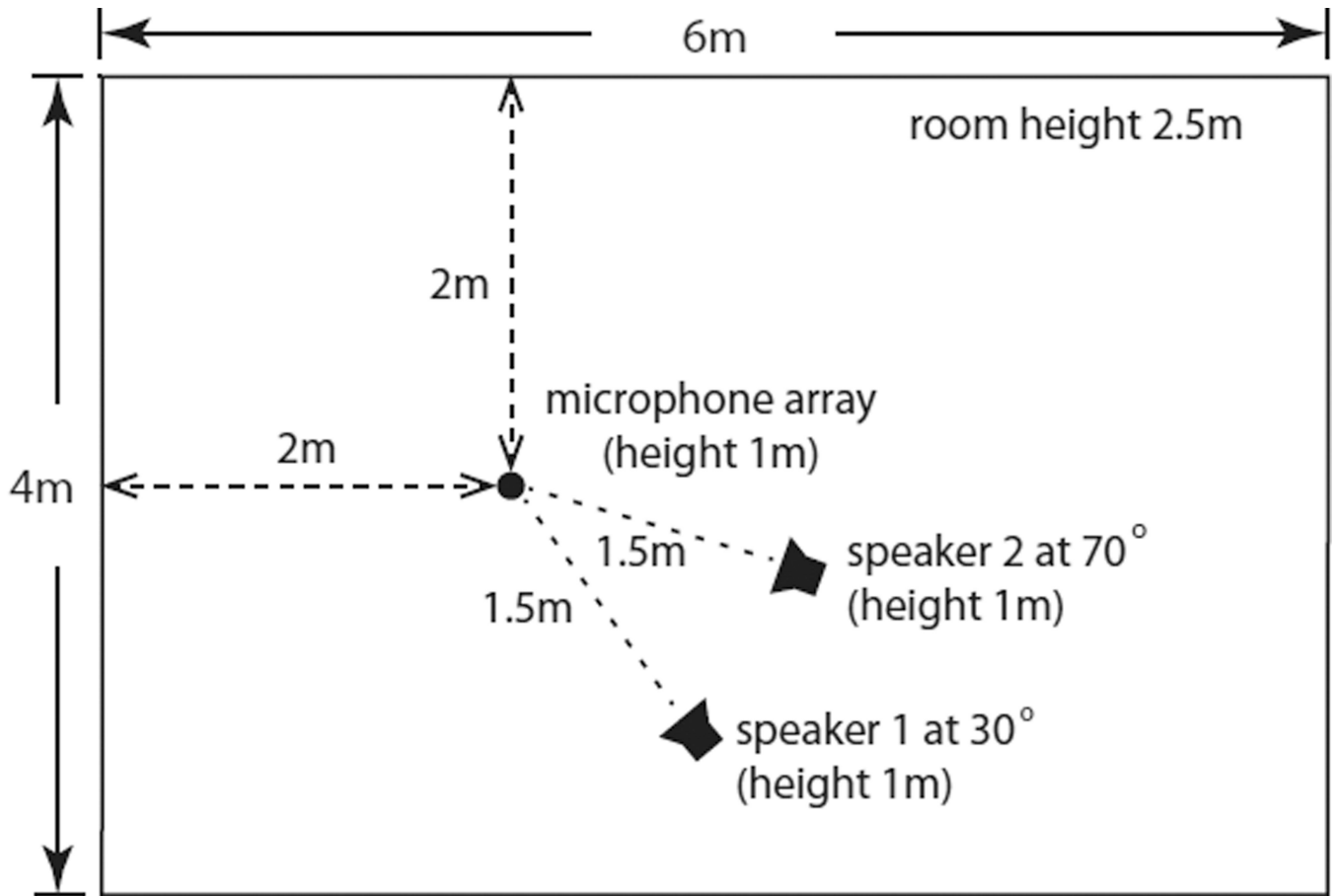


**Fig. 7.** SIR for separation of two sources incident on the microphone array. The azimuth angle of the first source is fixed at  $\theta_1=30^\circ$  and the azimuth angle of the second source  $\theta_2$  is varied from  $-15^\circ$  to  $135^\circ$ .

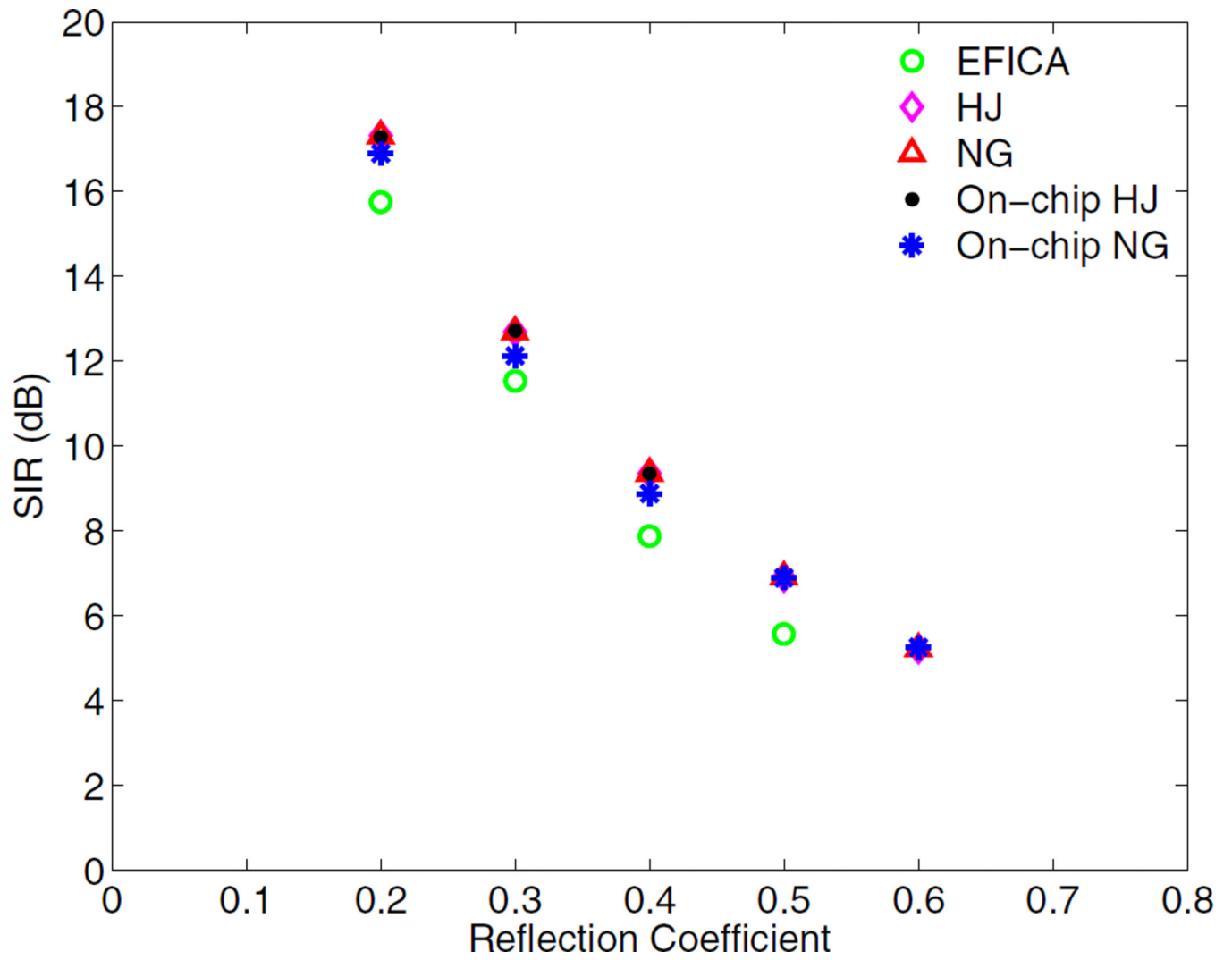


**Fig. 8.** Separation performance as a function of spatially uncorrelated sensor noise when the source signals are impinging the array at  $30^\circ$  and  $70^\circ$ .

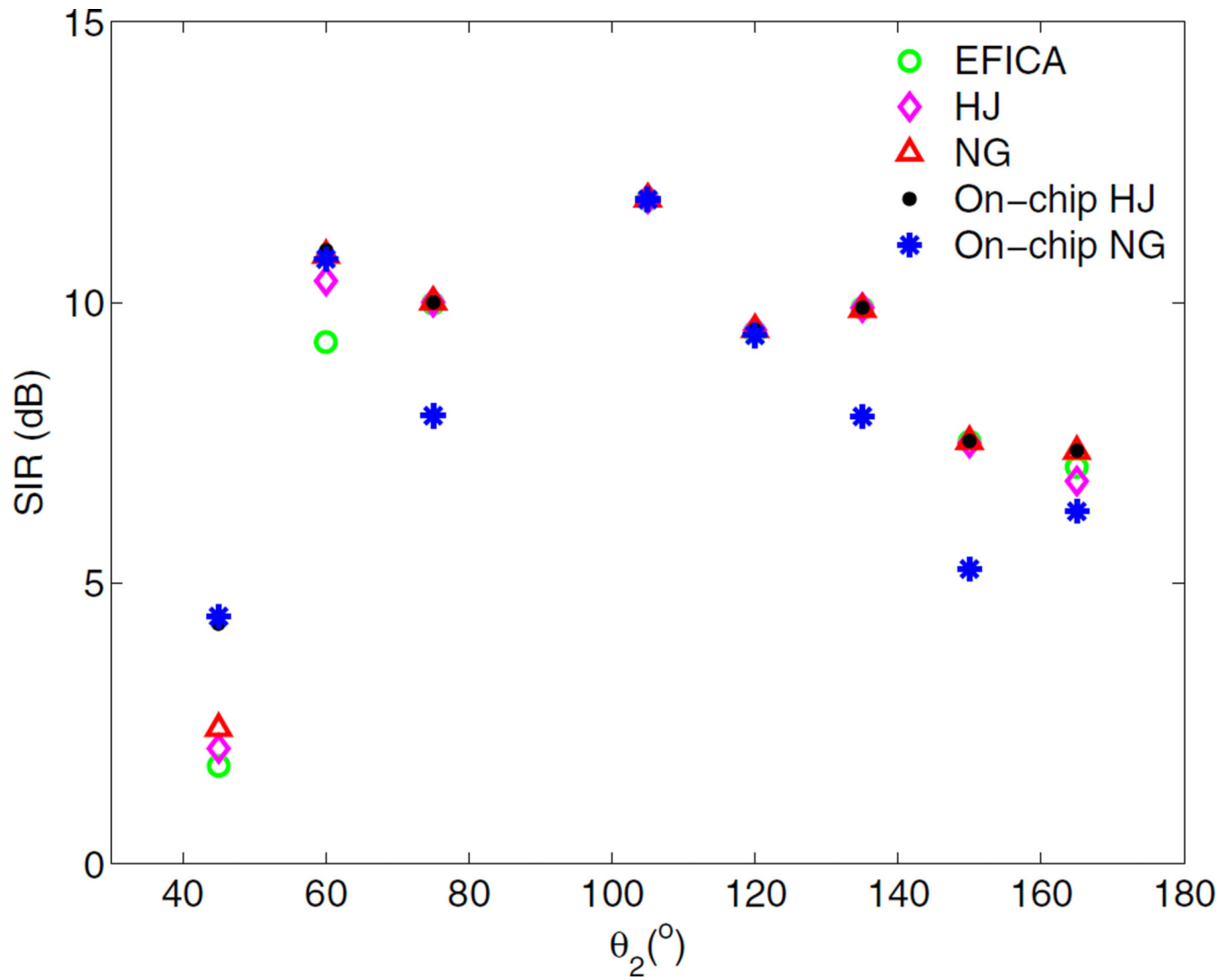




**Fig. 9.**  
Location of the sensor array and speakers in the simulated room environment.

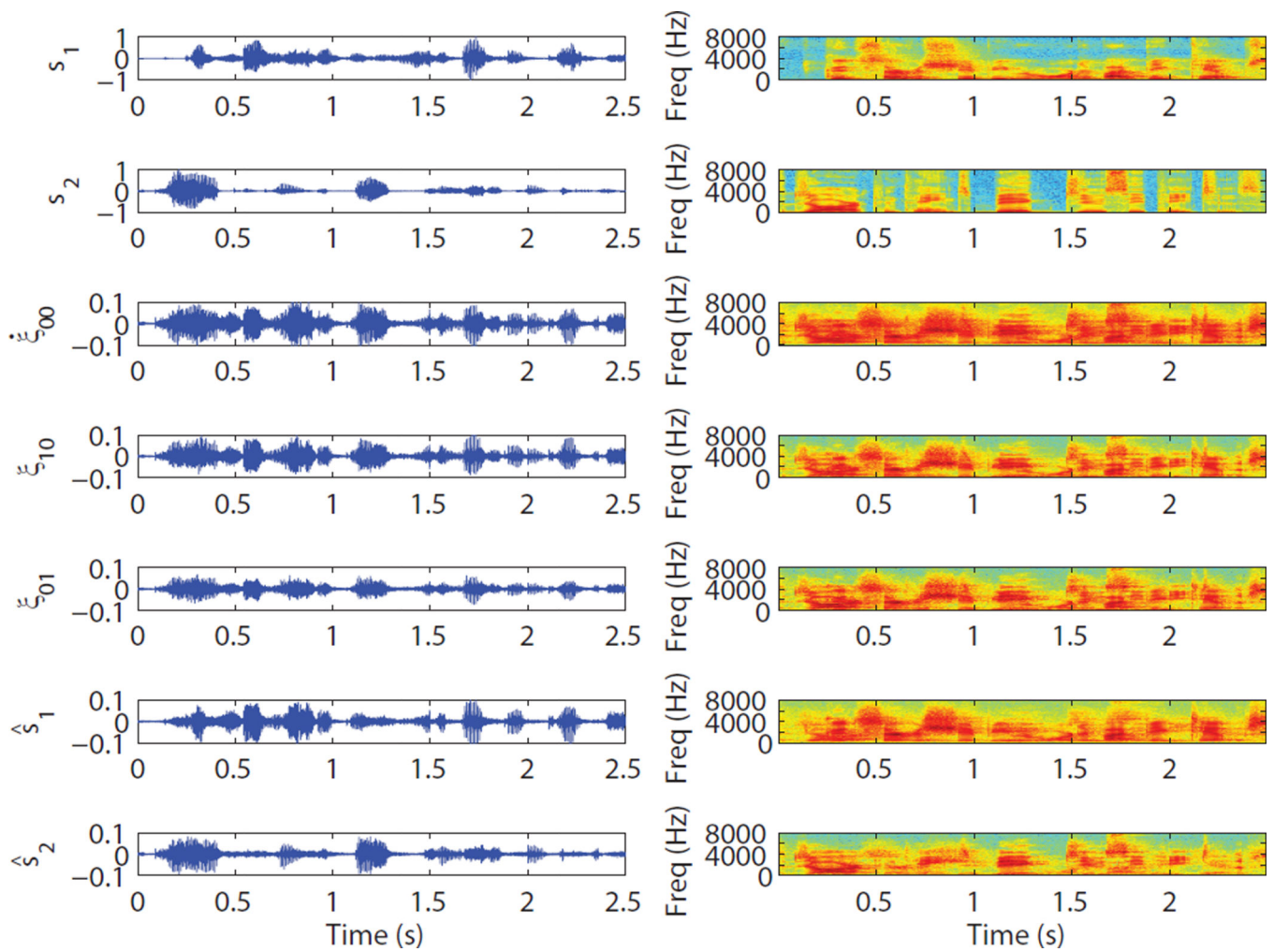


**Fig. 10.** Separation performance in reverberant environment as a function of reflection coefficient when the two sources are located at  $30^\circ$  and  $70^\circ$ .



**Fig. 11.**

SIR for two reconstructed speech sources played through a loudspeakers to a miniature microphone array in a conference room. The azimuth angle of the first source was fixed at  $\theta_1 = 30^\circ$  and the azimuth angle of the second source  $\theta_2$  was varied from  $45^\circ$  to  $135^\circ$ . The elevation angles  $\phi_1$  and  $\phi_2$  for both sources were  $8^\circ$ .



**Fig. 12.** Time waveforms and spectrograms of the source signals  $s_1$  and  $s_2$ , spatial gradients (3), the input signals to the ICA processor, and reconstructed source signals, output signals of the ICA processor for the case of two source signals with azimuth angles of  $\theta_1=30^\circ$  and  $\theta_2=105^\circ$ .

TABLE I

Comparison Results of Different ICA Implementations

	[7]	[8]	[9]	[10]	This work
Technology	FPGA	FPGA	90nm CMOS	90nm CMOS	0.5 $\mu$ m CMOS
Algorithm	Infomax ICA	FastICA	FastICA	FastICA	NG ICA
No of channels	2	2	8	8	3
Speed (MHz)	12.288	50	100	11	0.016
Power	98.8mW	NA	16.35mW	86 $\mu$ W	192 $\mu$ W

**TABLE II**

## ICA Processor Characteristics

Technology	0.5 $\mu\text{m}$ 2P3M CMOS
Size	3 mm $\times$ 3 mm
Supply	3 V
Room separation	6 dB – 12 dB
Power dissipation	192 $\mu\text{W}$ at 16 kHz

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript