Article

Sound Activity Monitor Circuit for Low Power Consumption of Always-On Microphone Applications

Jong Pal Kim

Advanced Research Center for Mechatronics Engineering, School of Mechatronics Engineering, Korea University of Technology and Education, Cheonan 31253, Republic of Korea; jongpalk@koreatech.ac.kr

Abstract: A novel sound activity monitor (SAM) circuit for low power consumption of always-on microphone applications is presented. To reduce average power consumption, the ultra-low-power SAM is essential and operates a readout integrated circuit (ROIC) in low power mode with silent input or in normal power mode with voice input. A novel SAM with an architecture that does not include an envelope detector is proposed to achieve low power consumption. A new architecture is also proposed to improve MEMS sensitivity by connecting the SAM input to the source follower (SF) output instead of connecting the SAM input to the MEMS port already connected to the SF. In addition, in order to prevent inefficient frequent operation mode conversion, a feature of delaying the transition to the low-power mode after the sound is silenced is implemented. The proposed architecture is designed and verified based on the standard 0.18 μm CMOS process. The SAM, which consists of two-stage amplifiers (OA, AMP2), comparators, and a logic circuit, consumes a 1 μA current. The analog path consisting of SF, OA, and AMP2 in low power mode has a maximum amplification gain of 63 dB and a noise of 72 nV$_{rms}$/$\sqrt{Hz}$ at 1 kHz.

Keywords: microphone; ROIC; low power; sound activity monitor (SAM)

1. Introduction

Recent advances in artificial intelligence based on deep learning technology have also made significant progress in speech recognition accuracy. According to the Internet Trends 2018 report, voice recognition accuracy reached 95% in 2017, which is the level of human recognition accuracy [1]. As voice recognition accuracy improves, voice recognition-based applications are also expanding to various applications such as artificial intelligence assistants, mobile payments, security identification, and translation [2]. As a result of this increased recognition accuracy and application expansion, the global voice recognition market is expected to reach $28.3 billion at a compound annual growth rate (CAGR) of 19.8% in 2026, according to Fortune Business Insights [3].

The microphone is used as a sensor for voice recognition and consists of a transducer and a readout integrated circuit (ROIC). The transducer converts sound pressure into a change in electrical characteristics, and the ROIC converts and amplifies electrical changes in the transducer into electrical signals. As transducers, electret condenser microphones (ECMs) were mainly used in 2010, but now MEMS microphones are increasingly being used [4]. The global microphone market is expected to reach $3.4 billion with a CAGR of 8.2% by 2027 [5]. Major microphone market share companies include Knowles Corporation (Itasca, IL, USA), AAC Technologies (Hong Kong, China), Goertek (Weifang, China), ST Microelectronics (Geneva, Switzerland), and Infineon Technologies (Neubiberg, Germany). Commercial microphones typically consume current from hundreds of μA to mA or more. In order to be used in intelligent microphone applications that can fulfill user instructions, sound must be able to be detected in a powered-on state at all times for 24 h. Actual voice instruction occurs intermittently during 24 h using the voice recognition function. Therefore, considering these application characteristics, the average power consumption of the microphone can be reduced.

As a traditional method of reducing power consumption, it can be approached from the viewpoint of reducing power consumption of major components. If the microphone ROIC is divided into two main parts, it is an analog signal processing part and an analog to digital converter (ADC). In order to reduce power consumption in the analog signal processing part, an approach of minimizing the current consumption of the sub-block circuit can be used. Alternatively, the average power consumption can be reduced by changing the amplifier's current consumption according to the required noise level [6,7]. Moreover, obviously, reducing the power consumption of the ADC is also important in reducing the power consumption of the microphone [8,9]. However, in terms of power consumption reduction, the high-level improvement approach may have a greater effect than the low-level improvement approach. A low-level approach means improving the sub-block circuit level, and a high-level approach means improving the top architecture or utilizing a use scenario. Therefore, in this paper, the use scenario is used as a strategy to reduce the power consumption of the microphone in the corresponding application.

Figure 1 shows a microphone operation scenario to achieve low average power consumption. The ROIC consists of a low drop-out voltage regulator (LDO) for generating power, a drive bias generator (DR) for driving the MEMS transducer, a main channel for measuring sound, and a sound activity monitor (SAM) for detecting the presence of sound. Figure 1a shows no sound input and the ROIC operating in a low power mode. In low power mode, only elements for the MEMS transducer drive and SAM functions operate normally, while others operate in low-power or power-off mode. Therefore, the SAM must be operated at extremely low power consumption. When the SAM detects an external sound, as illustrated in Figure 1b, the ROIC operates in a normal power mode. In normal power mode, the SAM detects silence and putsthe ROIC in a low power mode again.

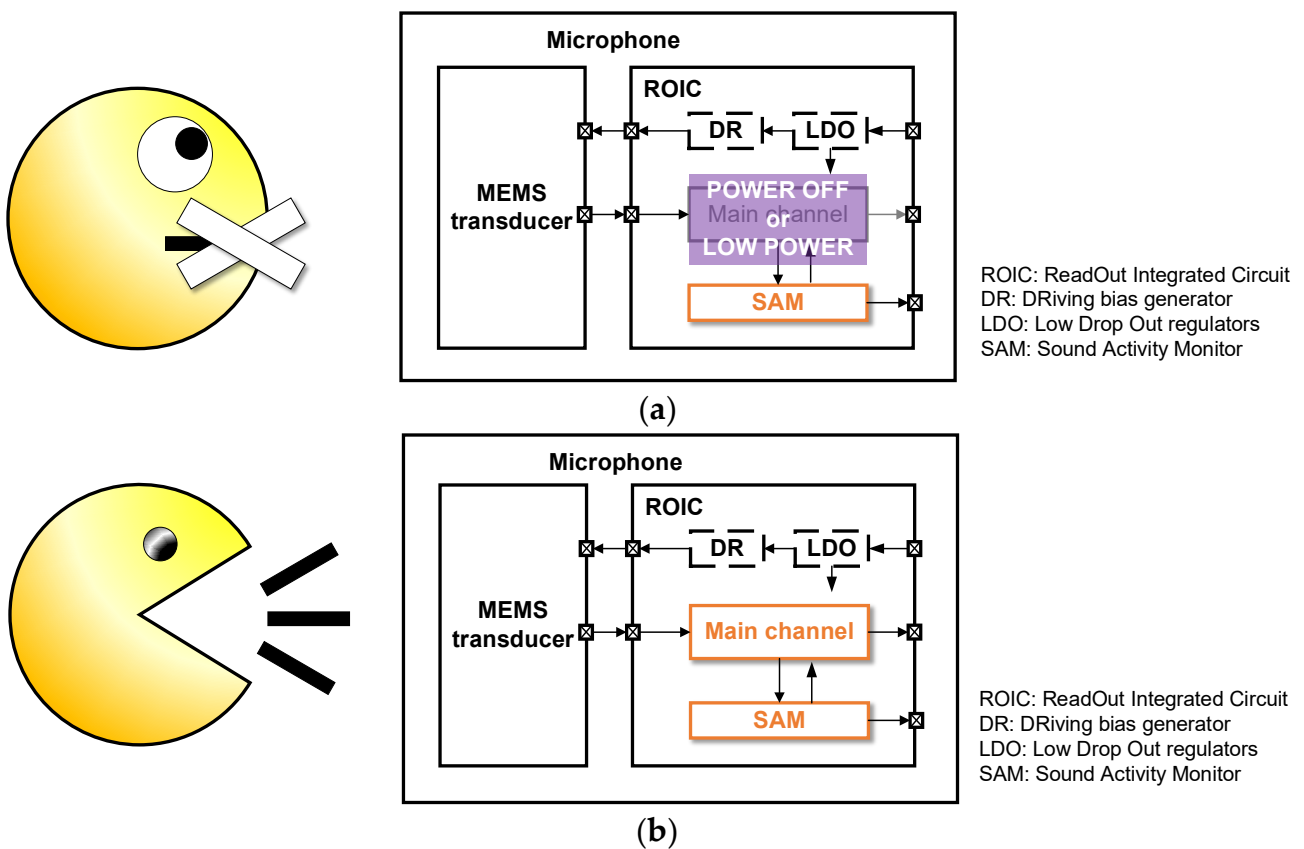

(a)

(b)

Figure 1. Microphone operating scenarios for low average power consumption: (**a**) Low power mode operation when there is no external sound activity; (**b**) Normal power mode operation when there is external sound activity.

Previous works on the SAM function of microphones are referenced in several papers [10–13]. In 2010, Tobi Delbruck showed an example of controlling the analog full channel as always on and only the digital part for speech analysis on/off [10]. Komail Bandami's paper published in 2016 showed an example of operating a basic analog measuring part in an always-on state and controlling the on/off control of an additionally mounted analog feature extractor [11]. In 2019, Minchang Cho showed an example of on/off control of a high-power channel, which is a parallel connection in terms of signal input, based on an ultra-low power (ULP) channel that is always on [12]. In the ULP channel, the bandwidth of the analog circuit is designed to be much lower than the full sound bandwidth for low power consumption. Instead, the high-frequency band should be downshifted to the base band by using the intermediate frequency to detect a sound in a frequency band higher than the circuit bandwidth. Therefore, in order to detect all sound frequency bands higher than the analog circuit bandwidth, a repetitive frequency down-shift operation is performed several times sequentially in time. Due to this time-sequentially repetitive operation, it takes up to several hundred ms to measure the desired full sound frequency band. In 2020, Youngtae Yang also presented an example of on/off control of the main channel with a low-power SAM [13]. In the previous work, the signal terminal of the MEMS transducer, the main channel input, and the SAM input were all connected, creating a structural problem of the parasitic capacitance of the SAM reducing the MEMS sensitivity. A MEMS transducer is being developed for miniaturization and unit price reduction. As a result, MEMS capacitance is also reduced and the effect of parasitic capacitance is increasing relatively. Therefore, an increase in parasitic capacitance due to the addition of SAM may be an important problem. A previous SAM consists of amplifiers, envelope detectors, comparators, and logic, and an attempt to simplify the architecture is required to achieve ultra-low power consumption. Therefore, in this paper, an improved technique for low power consumption of SAM and prevention of degradation of MEMS sensitivity due to the addition of SAM are presented.

2. Proposed Approach and Detailed Circuit

Figure 2 shows a top-view configuration of the circuit that will be covered in this paper. The top circuits include SF, SAM, and CLK_GEN. The SAM block consists of an open-loop amplifier (OA), a second amplifier (AMP2), a comparator (CMP), and a decision logic circuit (LOGIC). The CLK_GEN block generates the clock used by the LOGIC block. Initially, when there is no external sound input, the *WUP* signal is logically low (*LOW*) and the SF operates in low power mode. In low power mode, the source follower is not turned off completely and is operated with a low bias current of 100 nA level. The SF block consists of a stack type of PMOS acting as a source follower and NMOS supplying current to the source follower. The current flowing through the source follower can be controlled by adjusting the gate bias of the NMOS. When an external sound starts to be input, the SAM recognizes it and the *WUP* signal becomes logically high (*HIGH*), so the SF starts to detect the sound in normal power mode. In normal power mode, the source follower operates with a normal bias current of 7 μA level.

The output of SF is amplified by OA and AMP2, CMP generates a 1-bit digital signal that responds to sound above a certain level, and LOGIC outputs a *WUP* signal with a low or high logic value depending on the situation.

The architecture of Figure 2 incorporates the four proposed approaches. The first approach concerns where the SAM's input terminals should be connected. The second approach relates to whether the SAM should have an accurate amplification gain. The third approach concerns whether it is possible to remove the conventional envelope detector that was in use. The fourth approach relates to how to realize the falling edge of the *WUP* signal when it sounds and then disappears.
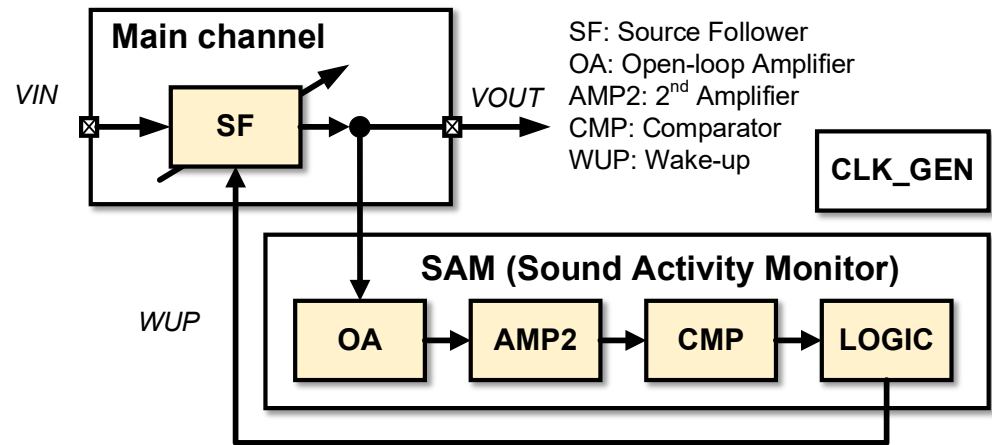
Figure 2. Top-level architecture of Sound Activity Monitor (SAM).

2.1. Approach 1: Connection Location of SAM Input

Figure 3 shows two cases of connecting SF and SAM. Figure 3a illustrates a case in which SF and SAM receive MEMS signals in common. This architecture has the advantage of being able to completely power off the SF in the absence of sound input, as SAM can directly monitor sound signals from the MEMS device. However, there is a problem of lowering the sensitivity of MEMS devices. The sensitivity of a MEMS device can be expressed by the following equation:

$$VIN = VDR \times C_m / [C_m + C_{p1} + C_{p2} + C_{p3}], \tag{1}$$

in which *VDR* is the driving voltage of the MEMS device, C_m is the MEMS capacitance of the microphone, C_{p1} is the parasitic capacitance of MEMS, C_{p2} is the input capacitance of SF, and C_{p3} is the input capacitance of SAM. As shown in Equation (1), the sensitivity of the MEMS device decreases due to the input parasitic capacitance C_{p3} of SAM. Therefore, in order to improve the MEMS sensitivity, it is necessary to minimize the parasitic capacitance connected to the input terminal from the MEMS. Figure 3b shows the proposed SAM connection configuration. The SAM is not connected to the input port of the MEMS, but to the output of the SF. In the proposed SAM connection configuration, the MEMS sensitivity is improved according to the following equation

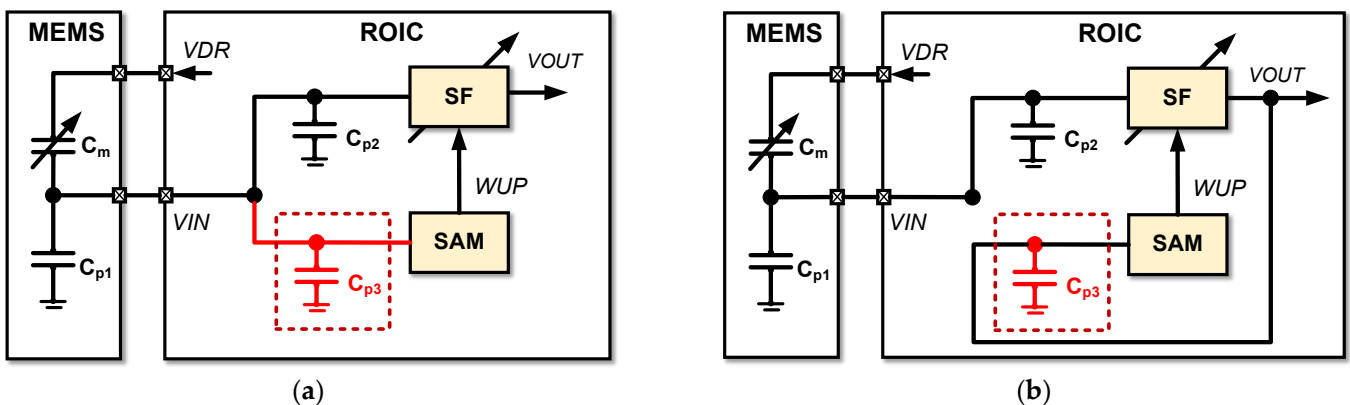$$VIN = VDR \times C_m / [C_m + C_{p1} + C_{p2}]. \tag{2}$$

(**a**)

(**b**)

Figure 3. Architecture approach 1: (**a**) Configuration when SF and SAM share MEMS input terminal; (**b**) Configuration when MEMS input connects only to SF.

When SF power is completely turned off in this configuration, the SAM cannot detect any sound signals. Therefore, SF should not be turned off completely even when there is no sound and should be operated to detect sound with minimal power consumption. For this reason, SF should be designed to have two operating modes: a normal power mode that can precisely measure sound and a low power mode that measures only the presence or absence of sound.

2.2. Approach 2: Front-End Amplifier Type on SAM

The front-end amplifier of the conventional SAM uses a closed-loop amplifier (CA) as shown in Figure 4a, and the front-end amplifier of the newly proposed SAM uses an open-loop amplifier (OA) as shown in Figure 4b. Typical amplifiers are designed in a closed-loop architecture to obtain accurate amplification gains. The front-end amplifier only needs to amplify a signal corresponding to a low sound by a certain amount or more, so an accurate amplification gain is not required. The closed-loop amplifier scheme to obtain an accurate gain consumes the area of four capacitors to create a gain ratio, a phase compensation capacitor, and a phase compensation resistor. The closed-loop amplifier should perform phase compensation to secure stability, which results in reduction of frequency bandwidth. Therefore, in order to obtain the desired closed-loop bandwidth, more current must be consumed to secure the higher bandwidth before compensation. In the case of open-loop amplifiers, capacitors for amplification ratio and stability compensation passive components are not required, improving area and power consumption.
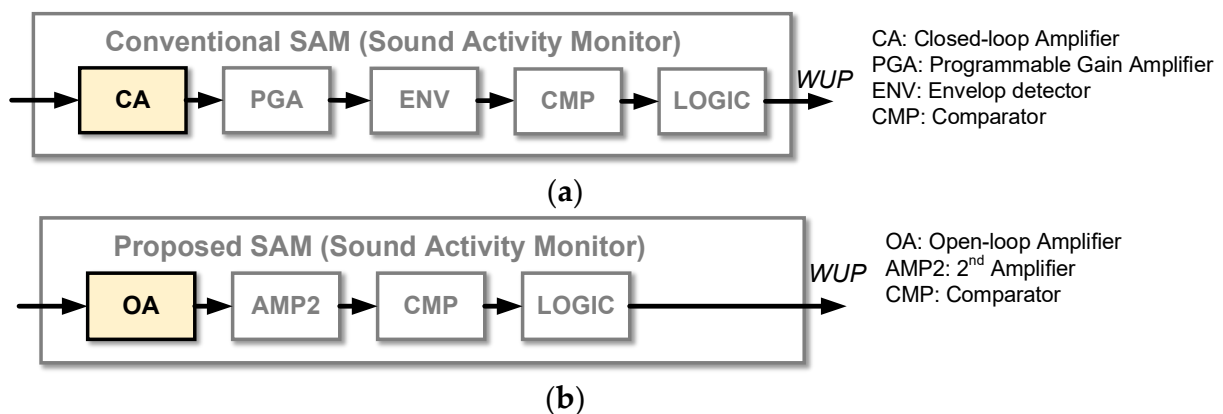

Figure 4. Architecture approach 2: (**a**) Closed-loop amplifier(CA) vs. (**b**) Open-loop amplifier(OA).

2.3. Approach 3: Removal of Envelope Detector

Figure 5a is a conventional SAM configuration. The envelope detector (ENV) extracts the envelope of the amplified signal via CA and PGA and compares the signal magnitude in the comparator (CMP). A typical envelope detector shown in Figure 5b consists of a big resistor (R_{env}) that causes a fine leakage current, a big capacitor (C_{env}) to maintain an output voltage, a current supply switch (SW_{env}) to increase a voltage of an output signal (OUT_{env}), and a comparator (CMP_{env}) that compares the amplitude of an input signal with the current output signal (OUT_{env}) [14,15]. If the magnitude of the signal (IN_{env}) is smaller than the magnitude of the output signal (OUT_{env}), the comparator output (CSW) becomes logically low (*LOW*) and the switch (SW_{env}) will be turned off. Then, as the charges stored in the capacitor (C_{env}) are discharged through the resistor (R_{env}), the voltage of the output signal (OUT_{env}) gradually decreases. As the output signal (OUT_{env}) gradually decreases and becomes smaller than the input signal (IN_{env}), the output of the comparator (CSW) becomes *HIGH* and the switch is turned on. While the switch (SW_{env}) is turned on, current is supplied through the switch (SW_{env}), so that a charge is supplied to the capacitor (C_{env}) and the output voltage (OUT_{env}) rises rapidly. Therefore, even if a DC signal is input to the envelope detector (ENV), the envelope detector output signal (OUT_{env}) takes a sawtooth

shape as charging and discharging are repeated in the capacitor (C_{env}). For this reason, as shown in Figure 5c, a ripple of size A_{ripple} occurs in the output signal of the conventional envelope detector. Since such a ripple (A_{ripple}) is not distinguished from the small sound signal, it defines a limitation of measurable sound. In addition, detecting the presence of sound after extracting the envelope of sound hinders rapid detection. Therefore, the use of envelope detectors in terms of low-sound detection and fast detection is disadvantageous, and the envelope detector is removed from the conventional SAM configuration as shown in Figure 4b.
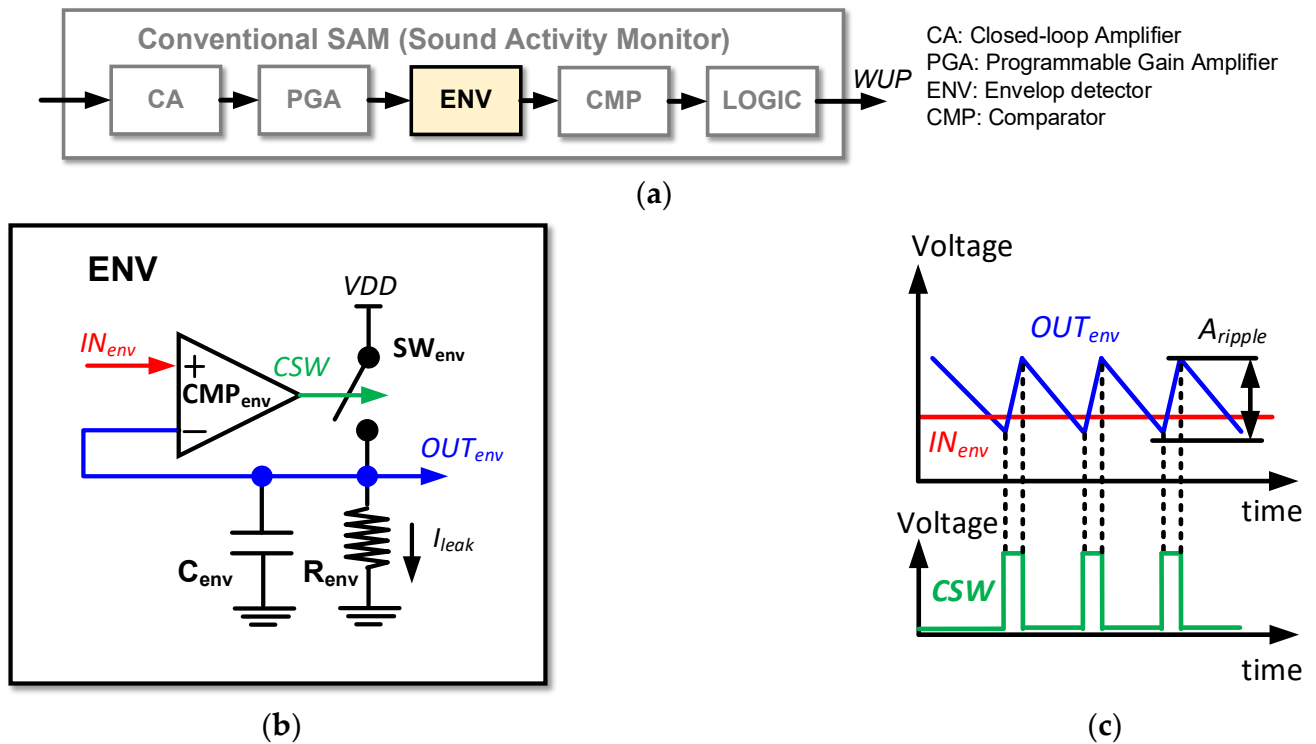

Figure 5. Architecture approach 3: (**a**) A conventional SAM includes a conventional envelope detector (ENV) and a comparator (CMP); (**b**) The configuration of a conventional envelope detection circuit (ENV); (**c**) Signal diagram of a conventional envelope detector.

2.4. Approach 4: Delay and Input-Blind Falling Edge of WUP Signal

It is necessary to determine under what conditions the wake-up signal (*WUP*) in the logical high state should be in the logical low state. Since the speech sound is in units of words and syllables, the sound may not be continuously formed. The wake-up signal (*WUP*) should not be made *LOW* as soon as the sound disappears because the sound may be cut off for a while and then continue again. Therefore, it is necessary to wait and see whether the silence continues for a predetermined time (t_{fd}) even without sound input, as shown in Figure 6. The time delay (t_{fd}) before the falling edge shall be adjustable. When a falling edge occurs in the *WUP* signal, the SF starts operating in low power mode. At the falling edge of the *WUP* signal, a glitch occurs in the SF output and is mistaken for a sound signal, allowing the *WUP* signal to immediately return to *HIGH*. Therefore, in order to prevent malfunction at the *WUP* falling edge, it is necessary to prevent the rising edge of the *WUP* signal for a time (T_{fb}) long enough for the glitch to disappear after the *WUP* falling edge. To prevent the *WUP* from returning to *HIGH* immediately after the falling edge, the time delay (t_{fb}) of more than half clock is sufficient.
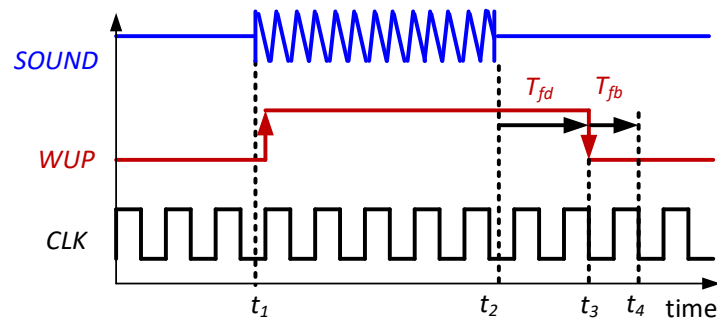
Figure 6. Architecture approach 5: After a predetermined delay time (T_{fd}) elapses after the sound input disappears at t_2, a falling edge at t_3 is expressed as a wakeup signal (*WUP*).

2.5. Detail Circuits

Figure 7 shows a detailed circuit diagram of the source follower block (SF). The device that functions as a source follower is MP_{10} and transmits an AC signal from the input (*VIN*) to the output (*VOUT*). The DC voltage on the *VIN* is set to V_b via high resistance (R_{bsf}). The high-resistance value of the resistor (R_{bsf}) is implemented as a pseudo-resistor using a PMOS. In the source follower (MP_{10}), current flows as much as I_{LP} in the low power mode and as much as ($I_{LP} + I_{NP}$) in the normal power mode. When the *WUP* is *LOW*, it becomes low power mode, and the current I_{LP} mirrored by MN_{12} and MN_{10} flows through the source follower MP_{10}. Since the switch SW_{np} is turned off, the switch SW_{lp} is turned on, and the MOS MN_{11} is turned off, no additional current is added to the source follower MP_{10}. When the *WUP* is *HIGH*, it becomes normal power mode, and the current I_{LP} mirrored by MN_{12} and MN_{10} is added to the current flowing through the source follower MP_{10}. In addition, since the switch SW_{np} is turned on and the switch SW_{lp} is turned off, current I_{NP} mirrored through MN_{13} and MN_{11} is added to the source follower MP_{10}. Eventually, in normal power mode, a current of ($I_{LP} + I_{NP}$) flows through the MOS MP_{10}.
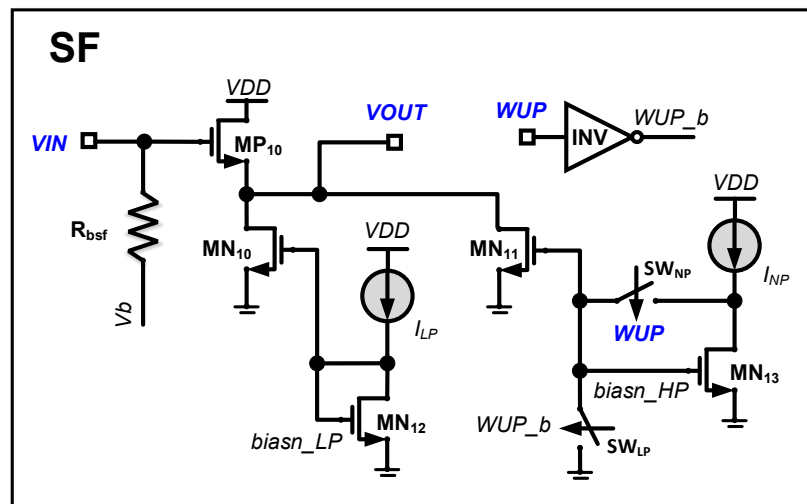

Figure 7. Detailed circuit diagram of the source follower stage.

Figure 8 shows the detailed circuit of the OA block. The OA consists of a high-pass filter block (HPF), an amplification block (Ainv), and a common-mode feedback block (CMFB). The high-pass filter (HPF) removes DC and low-frequency components input from SF and is implemented with MIM capacitor (C_{20}) and PMOS-based pseudo-resistor (R_{oa}). The amplification block (Ainv) is implemented as an inverter-based amplification architecture. Inverter-based amplifiers are commonly used because of their simple structure and good efficiency [16,17]. While inverter-based amplifiers may have poor linearity for

large inputs, nonlinearity during amplification is not an issue in SAM applications. Current I_{oa} is supplied from the top of Ainv and flows through a differential inverter-based amplifier to NMOS (MN$_{21}$) functioning as common-mode feedback control. When the common-mode voltage of OA output is greater than $VDD/2$, $Vcmfb$ increases and the current flow of NMOS improves, so that the common-mode voltage of OA output decreases. On the contrary, when the common-mode voltage of the OA output is less than the $VDD/2$, the $Vcmfb$ decreases and the current flow of the NMOS (MN$_{21}$) decreases, thereby increasing the common-mode voltage of the OA output.
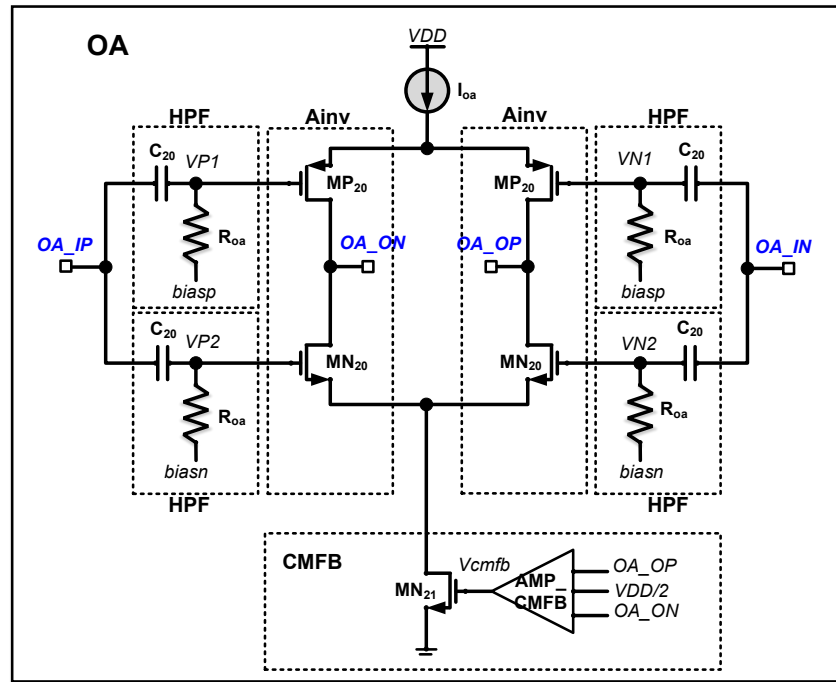

Figure 8. Detailed circuit diagram of the open-loop amplifier (OA).

Figure 9 shows the detailed circuit of the AMP2 block. The AMP2 performs additional amplification, low cutoff-frequency adjustment, and common-mode signal removal. AMP2 has a charge amplifier structure, and the amplification ratio is defined as a ratio of C_{in}/C_f [18–20]. The low cutoff frequency (f_{c_low}) is determined by the feedback resistor R_f and the feedback capacitor C_f. as in Equation (3). The low cutoff frequency can be changed by adjusting the feedback resistance. Since the feedback resistor R_f needs to have a large resistance value, it is implemented as a PMOS-based pseudo-resistor. The internal amplifier (AMP_AMP2) of the second amplifier (AMP2) consists of a two-stage amplifier structure and a common-mode feedback block.

$$f_{c\_low} = 1/[2\pi\, R_f \times C_f] \tag{3}$$

Figure 10 shows the detailed circuit of the CMP block. Differential input is amplified using a well-known positive feedback architecture [21]. The operation of the comparator when the value of CMP_IN momentarily decreases while in a complete symmetrical state will be described as follows. A momentary decrease in the gate voltage CMP_IN of MP$_{40}$ causes an instantaneous increase in the current i_{p40}. A current i_{p40} flowing in the MP$_{40}$ is composed of a current i_{n40} flowing in the MN$_{40}$ and a current i_{n42} flowing in the MN$_{42}$. Since the gate voltage of MN$_{42}$ is still fixed to $Vg43$, MN$_{42}$ current i_{n42} flows as before, and an increase in current i_{p40} leads to an increase in current i_{n40}. As the $Vg40$ increases, the current i_{n41} of MN$_{41}$ increases and the current i_{n43} of MN$_{43}$ decreases. As the current i_{n43} decreases, the voltage $Vg43$ also decreases, and the current i_{n42} of MN$_{42}$ also decreases. Due to the increase in i_{p40}, i_{n42} decreases and i_{n40} increases continuously. As a result, $Vg40$

continuously increases and *Vg43* continuously decreases. The increased *Vg40* increases i_{p42}, decreases *Vg42*, and increases the current flowing through MP$_{43}$. Meanwhile, the reduced *Vg43* reduces the current flowing through the MN$_{45}$. The increase in the current of MP$_{43}$ and the decrease in the current of MN$_{45}$ cause *CMP_OUT* to become *HIGH*. In the same principle, if *CMP_IN* is higher than *CMP_IP*, the output *CMP_OUT* becomes a logical low state.
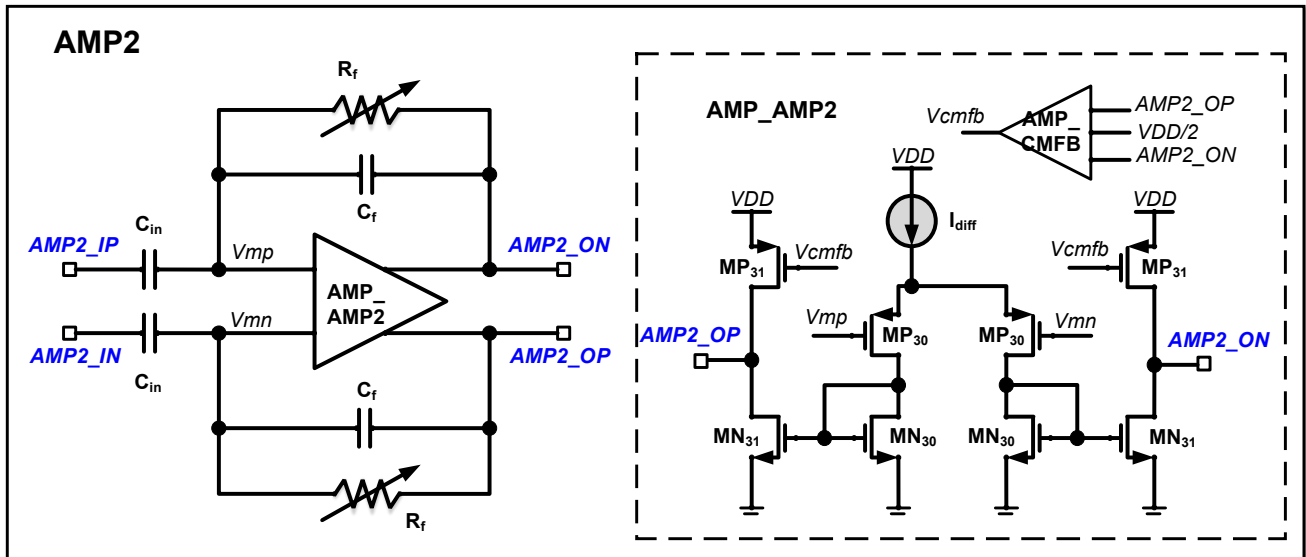

Figure 9. Detailed circuit diagram of the second amplifier (AMP2).
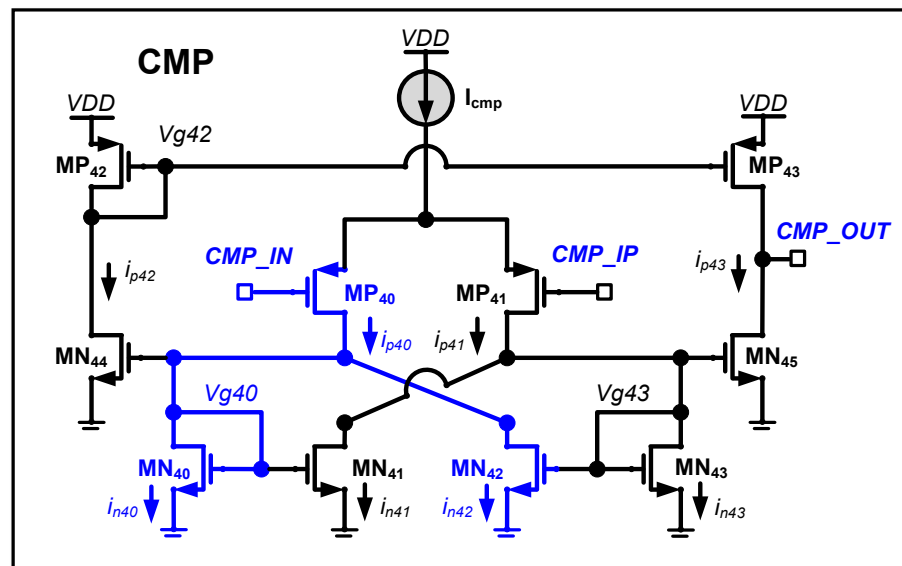

Figure 10. Detailed circuit diagram of the comparator (CMP).

Figure 11 shows the workflow and detailed circuit diagram of LOGIC, and Figure 12 shows the main signal clock diagram of LOGIC.

In Figure 11a, at step F1, the LOGIC continuously monitors whether a rising edge or a falling edge appears in the comparator output according to the sound input. When a rising edge or falling edge is detected in the output of the comparator, the *WUP* rising edge is output in step F2. In step F3, it is monitored whether there is no sound input for a predetermined time (T_{fd}). If the silence lasts for a predetermined time (T_{fd}), a falling edge is

output to the *WUP* signal in step F4. After a predetermined delay (T_{fb}) in step F5, the logic starts to recognize the comparator output again in step F6.

Figure 11b shows the detailed circuit configuration of the LOGIC. The LOGIC consists of four blocks: 'A. WUP_RISING_GEN' and 'B. WUP_OUTPUT', 'C. WUP_FALLING_GEN' and 'D. EN_INPUT_GEN'. In block "A. WUP_RISING_GEN", after the LOGIC reset, both the output *INA* of the D flip-flop DFFA and the output *INB* of the D flip-flop DFFB are in the low state. When a rising edge occurs in the input *IN*, the DFFA outputs *HIGH* to the output *INA*, and when a falling edge occurs in the input *IN*, the DFFB outputs a logic low value to the output *INB*. The DFFA output *INA* and DFFB output *INB* use a OR gate to generate a *CMP_edge* signal. That is, whether a rising edge or a falling edge occurs in the input signal *IN*, a rising edge at t_1 is generated in *CMP_edge* by a faster edge at *IN* as shown in Figure 12.

In block 'B. WUP_OUTPUT', the rising edge of the *CMP_edge* is transferred from block A, the output *WUP* is changed from *LOW* to *HIGH* at time t_1, and the output *WUP* is changed from *LOW* to *HIGH* at time t_3 by receiving a low active reset signal of *WUP_FALLING* from block C.
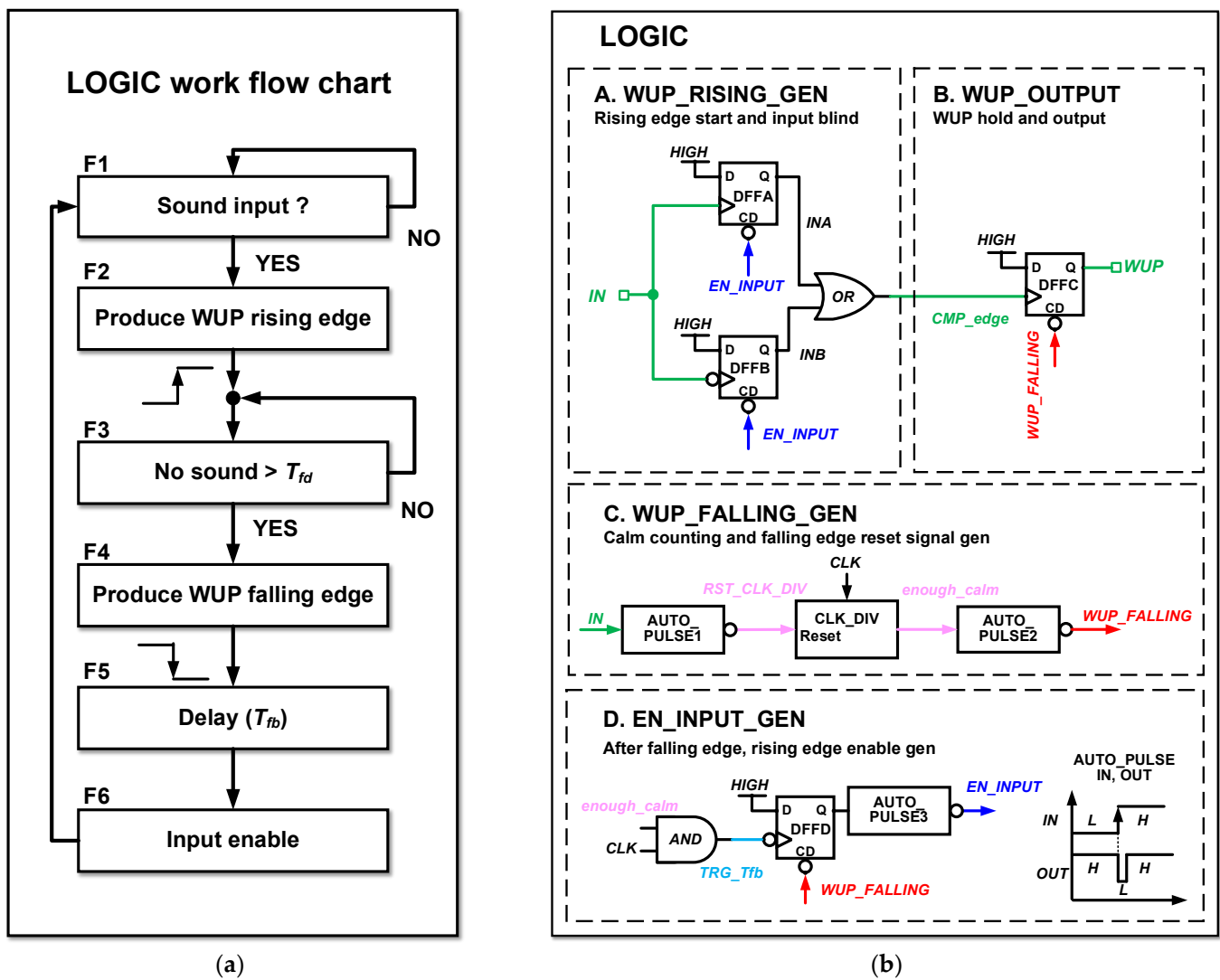

Figure 11. Workflow and circuit of LOGIC: (**a**) Working flow chart; (**b**) Detailed circuit diagram.
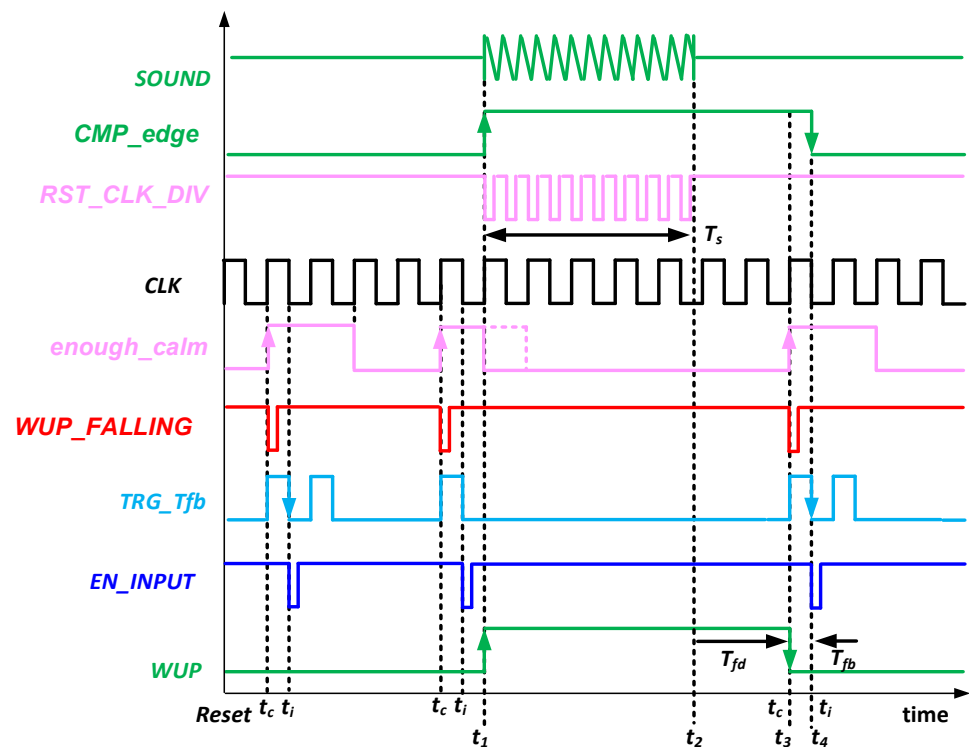
Figure 12. Clock diagram showing the operating principle of LOGIC.

The *WUP* signal with *HIGH* should be changed to a logical low value if the silence persists. Block 'C. WUP_FALLING_GEN' generates a low active reset signal of *WUP_FALLING* at t_3 when there should be a falling edge of *WUP*. Block C consists of CLK_DIV and two AUTO_PULSEs before and after CLK_DIV. As shown in the lower right corner of Figure 11b, the AUTO_PULSE typically outputs *HIGH* and then generates a low active reset signal as soon as a rising edge occurs on the next input signal. The block CLK_DIV divides the frequency of *CLK* by a multiplication ratio of 2 according to the register setting. As can be seen in the T_s interval in Figure 12, if the sound input continues to generate rising edges at the input *IN*, AUTO_PULSE1 generates a repeated low active reset signal on the output signal *RST_CLK_DIV* and resets the CLK_DIV. Then, the CLK_DIV output (*enough_calm*) remains in the low state continuously. At every t_c in Figure 12, the rising edge of *enough_calm* generates a low active reset signal in the output signal *WUP_FALLING* of AUTO_PULSE2. Eventually, the low active reset signal of *WUP_FALLING* resets the DFFC, and a falling edge is generated in the *WUP* signal. As a result, the time T_{fd} from the start of silence to the occurrence of the falling edge of the *WUP* corresponds to the half-cycle time of the signal *enough_calm*. If there is continuous silence, *enough_calm* has a periodic square waveform, and a periodic low active reset signal is generated in *WUP_FALLING* to keep the *WUP* signal at a low logical value as shown in Figure 12.

Block 'D. EN_INPUT_GEN' generates a reset signal *EN_INPUT* and provides it to block A, allowing block A to detect the rising edge or falling edge of *IN* again after time T_{fb} has elapsed from the falling edge of the *WUP*. When a falling edge occurs at t_3 in the *WUP*, the SF switches to the low power mode. At this time, the glitch generated in SF causes an edge change as if a sound was input to LOGIC input *IN*. In order to avoid responding to these fake sound signals, it is necessary to delay the start of the input detection function of block A after the falling edge occurs in the *WUP*. While the silence is maintained at the beginning after reset, the output of D flip-flop (DFFD) has *HIGH*, and the output (*EN_INPUT*) of AUTO_PULSE3 also has *HIGH*. When the sound starts at t_1 and stops at t_2, low active reset occurs in *WUP_FALLING* at t_3, and output terminal Q of DFFD becomes *LOW*. To generate the *TRG_Tfb* signal, *enough_calm* and *CLK* are logically calculated using

an AND gate. As can be seen in Figure 12, the *TRG_Tfb* signal shows a waveform in which the *CLK* signal appears only when the *enough_calm* is in *HIGH*. After DFFD is reset by *WUP_FALLING* at each t_c, a low active reset signal at t_i is generated in *EN_INPUT* at the first falling edge of *TRG_Tfb*. After t_{fb} time from the falling edge of the *WUP*, that is, after the half cycle time of *CLK*, block A is reset. The initialized A block can deliver the edge change of the *IN* signal again to the *CMP_edge* signal.

Figure 13 shows the detailed circuit diagram of the clock generator block (CLK_GEN) that generates the clock (*CLK*) used in LOGIC. The CLK_GEN block has a structure in which three clock delay elements (THY) are connected in series. The THY block was referred to a delay element based on thyristors [22]. Clock *INP* and clock *INN* have opposite phases, and clock *Q* and clock *Qb* also have opposite phases. The THY block is composed of four inverters and one capacitor. The first inverter (INV1) consists of MP_{111} and MN_{111}, the second inverter (INV2) consists of MP_{112} and MN_{112}, the third inverter (INV3) consists of MP_{113} and MN_{113}, and the fourth inverter (INV4) consists of MP_{114} and MN_{114}.
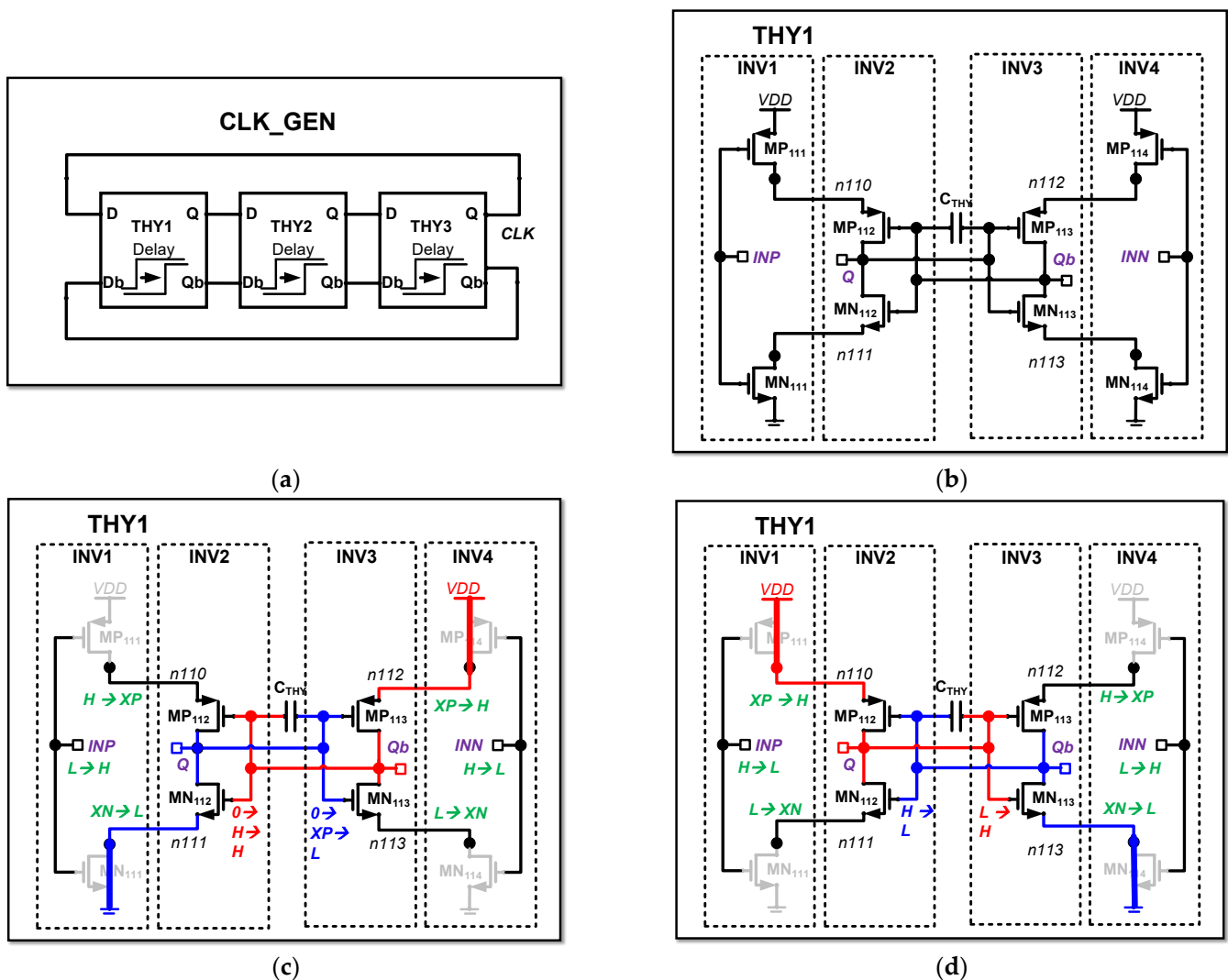

Figure 13. Detailed circuit diagram of CLK_GEN: (**a**) Configuration of CLK_GEN; (**b**) Detailed circuit of THY; (**c**) Operation voltage when INP becomes *HIGH*; (**d**) Operation voltage when INP becomes *LOW*.

Figure 13c shows the change in operating voltage of each node when *INP* changes from *LOW* to *HIGH* and *INN* changes from *HIGH* to *LOW*. Initially, when the *INP* of INV1 changes from *LOW* to *HIGH*, PMOS MP_{111} turns off, the voltage at node *n110*

gradually decreases to *XP*, NMOS MN_{111} turns on, and the voltage at node *n111* goes *LOW* immediately. When the *INN* of INV4 changes from *HIGH* to *LOW*, PMOS MP_{114} turns on, the voltage at node *n112* immediately goes *HIGH*, NMOS MN_{114} turns off, and the voltage of node *n113* becomes *XN*. Assuming that the gate voltages of INV2 and INV3 were both zero at the beginning, both PMOS MP_{112} and MP_{113} are turned on to induce the voltage of *XP* in *Q* and the voltage of *HIGH* in *Qb*. The *HIGH* voltage of node *Q* is higher than the *XP* voltage of node *Qb*, so NMOS NM_{112} turns on stronger than MN_{113} and *Q* becomes *LOW*. When *Q* becomes *LOW*, PMOS MP_{113} is completely turned on, *Qb* becomes strong *HIGH*, and *Qb* turns on NMOS MN_{112} completely to keep *Q LOW*.

Figure 13d shows the change in operating voltage of each node when *INP* changes from *HIGH* to *LOW* and *INN* changes from *LOW* to *HIGH*. When the *INP* of INV1 changes from *HIGH* to *LOW*, PMOS MP_{111} turns on, the voltage at node n110 goes *HIGH* immediately, NMOS MN_{111} turns off, and the voltage at node *n111* becomes *XN*. When the *INN* of INV4 changes from *LOW* to *HIGH*, PMOS MP_{114} turns off, the voltage at node *n112* goes *XP*, NMOS MN_{114} turns on, and the voltage of node *n113* becomes strong *LOW*. In the previous state, the gate voltage of INV2 was *HIGH* and the gate voltage of INV3 was *LOW*. The gate voltages of INV2 and INV3 are not in a strong state, but are formed by the charge stored in the capacitor C_{THY}. A leak occurs in the turned-off PMOS MP_{112} and NMOS MN_{113}, causing the voltages of *Q* and *Qb* to slightly increase or decrease, respectively. At some point over time, the voltage on *Q* decreases sharply, and the voltage on *Qb* increases sharply. Eventually, the voltage at *Q* becomes *LOW* and the voltage at *Qb* becomes *HIGH*. When *Qb* goes *LOW*, the PMOS MP_{112} is fully turned on and *Q* becomes strongly *HIGH*. A strong *HIGH* of *Q* turns the NMOS MN_{113} on completely, and *Qb* becomes a strong *LOW*. A *Qb* that becomes a strong *LOW* holds *Q* again a strong *HIGH*.

3. Simulation Verification and Results

Figure 14 shows the layout of the proposed circuit and the regions of the main blocks of SF and SAM. The layout was performed based on the 0.18 μm CMOS process, and the size of the circuit core was 0.71 mm × 0.39 mm.
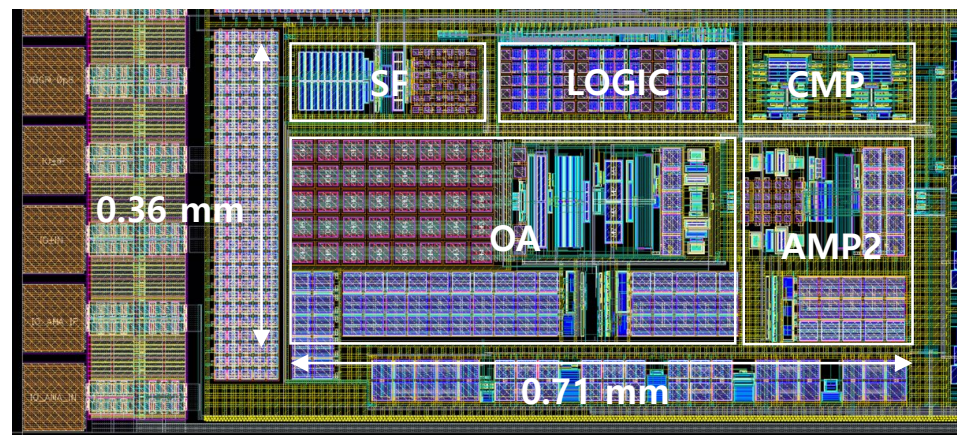

Figure 14. The layout of the proposed circuit and the regions of the main blocks of SF and SAM.

Figure 15 shows the frequency response and noise characteristics of analog blocks. In Figure 15a, the dotted line and the solid line correspond to the frequency response characteristics when SF is in the low power mode and the normal power mode, respectively. In the low-power mode, a current of 100 nA flows through the input MOS of the source follower and has a bandwidth of about 56 kHz. In the normal power mode, a current of 7.6 μA flows through the input MOS of the source follower and has a bandwidth of about 3.5 MHz.
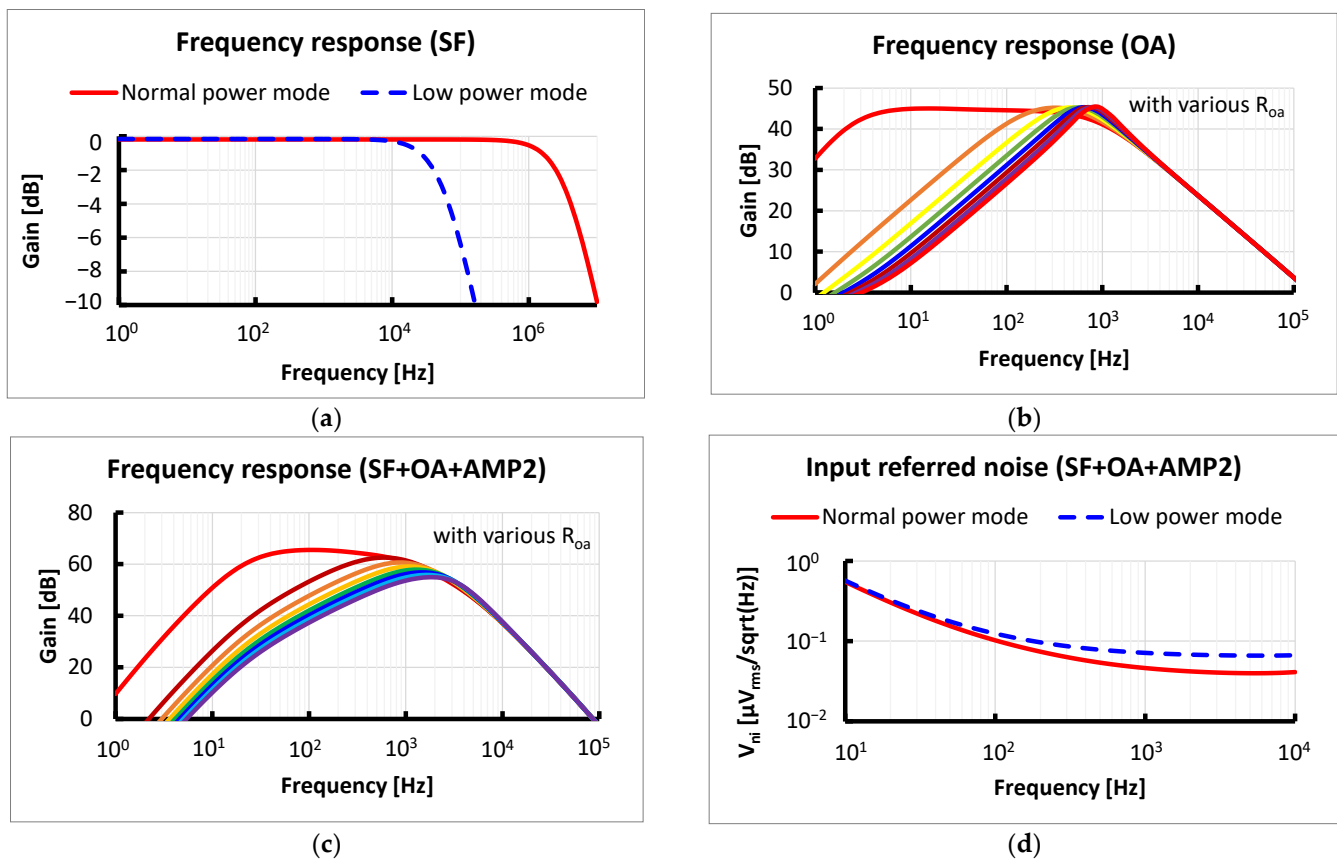
Figure 15. Characteristics of analog blocks: (**a**) Frequency response characteristics of SF; (**b**) Frequency response characteristics of OA; (**c**) Frequency response characteristics of SF + OA + AMP2; (**d**) Noise characteristics of SF + OA + AMP2.

Figure 15b shows the frequency response characteristics of the OA block. The high-pass cutoff frequency can be adjusted by changing the resistance R_{oa} value of the high-pass filter (HPF) of the OA block input part. The maximum amplification gain is 45.5 dB (172 V/V), and the low cutoff frequency can be adjusted in eight steps within the range of 2.9 Hz to 540 Hz. The OA block is a single-stage amplifier that directly utilizes an open-loop gain of 170 V/V for signal amplification. There are no stability issues as it does not use closed-loop feedback to get the correct gain, and there is no saturation in the absence of an input.

Figure 15c shows frequency response characteristics of blocks continuously connected in the order of SF, OA, and AMP2. It can be seen that the low cutoff frequency varies from 30 Hz to 720 Hz according to the AMP2 feedback resistance setting value. The high cutoff frequency is also affected by the AMP2 feedback resistance value Rf setting and ranges from 1.4 kHz to 3.7 kHz. The amplification gain has a maximum of 63 dB (1395 V/V) and a minimum of 55 dB (560 V/V). Because 'OA + AMP2' blocks have much lower bandwidth than the SF bandwidth of 56 kHz in low-power mode, their combined frequency response characteristics show the same characteristics regardless of the power mode of SF.

Figure 15d shows the noise characteristics of the combined blocks of SF, OA and AMP2. The SF noise characteristics vary depending on the power mode, and for this reason the noise characteristics of the 'SF + OA + AMP2' combined block vary depending on the power mode. The dotted line shows the noise characteristics when SF is in low power mode, and the solid line shows the noise characteristics when SF is in normal power mode.

Comparing the noise levels of the 'SF + OA + AMP2' block based on the frequency of 1 kHz, the noise floor value has 72 nV$_{rms}$/$\sqrt{Hz}$ in low power mode and 40 nV$_{rms}$/$\sqrt{Hz}$

in normal power mode. The integrated noise for the range from 600 Hz to 2 kHz has a 2.6 μV_{rms} in low power mode and a 1.7 μV_{rms} in normal power mode.

Due to the SAM connection architecture in this paper, the sensitivity of the MEMS transducer itself is the same whether or not SAM is added. Therefore, the overall microphone noise level for each power mode is determined by the noise level of the ROIC. Compared to the noise level of 1.7 μV_{rms} in normal power mode, the noise level in low power mode is 2.6 μV_{rms}, so the noise level is about 50% higher. Numerically, therefore, the overall noise performance of the microphone is 50% higher in low power mode than in normal power mode. However, the meaning of noise level in normal power mode and low power mode can be slightly different. In normal power mode, low noise means whether small changes in volume can be measured when there is sound. In low power mode, if a command is entered in silent state, the user is intentionally trying to enter a voice command and therefore does not speak quietly. Therefore, it may not be a big problem that the noise of the low power mode is about 50% greater than the noise of the normal power mode. This is especially true because when a sound is detected, it immediately switches to normal power mode before a single syllable has passed and takes measurements at a lower noise level. However, further research is needed on how appropriate the noise level is in low power mode.

Figure 16 shows the simulation results for the full chain from SF to LOGIC. As shown in the *SF_INPUT* of Figure 15, a signal consisting of no wave for 0.5 s and a sine wave having an amplitude of 1 mV_{pk} and a frequency of 1 kHz for 0.5 s was repeatedly supplied to the SF. As soon as the signal was input at 0.5 s, a rising edge was generated at the *CMP_edge*, which soon generated a rising edge of the *WUP* signal. The input signal was cut off at 1 s, but it can be seen that the *WUP* signal continues to maintain at *HIGH* even after 1 s. After the *SF_INPUT* signal was cut off at 1 s, a rising edge occurred in the *enough_calm* signal after 2 cycles of *CLK*, and the low active reset signal in the *WUP_falling* signal caused a falling edge in the *WUP* signal. Since the *CMP_edge* is still in *HIGH*, the input reaction of the LOGIC block "A. WUP_RISING_GEN" is in the deactivated state. Afterwards, it can be seen that the low active reset signal of the *EN_INPUT* signal is generated at the first falling edge of the *TRG_Tfb* signal generated by synthesizing *enough_calm* and *CLK*. The low active reset signal value of the *EN_INPUT* signal causes the *CMP_edge* signal to become *LOW*, and the block A of the LOGIC can react to the input again. Accordingly, it can be seen that the *WUP* signal becomes *HIGH* again as the sine wave is input to the *SF_INPUT* at 1.5 s.
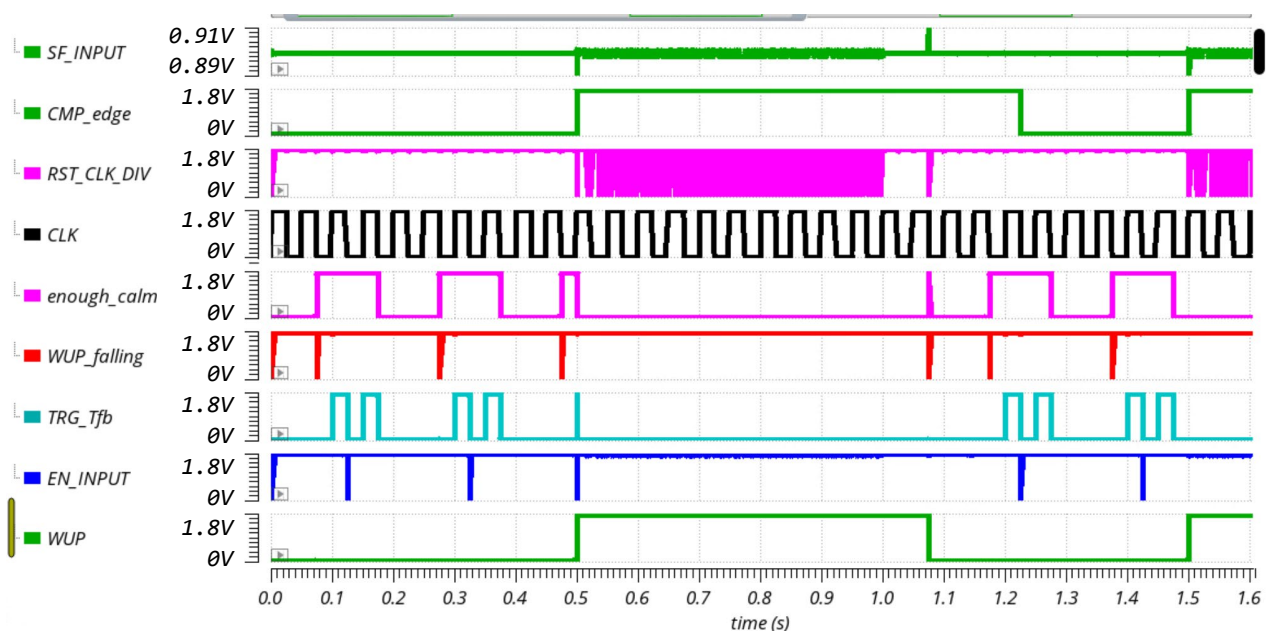

Figure 16. Full channel simulation results from SF to LOGIC.

The designed circuit operates at 1.8 V, and the SF block consumes a current of 200 nA in low power mode and a current of 14.3 μA in normal power mode. The details of the DC current consumption of the SAM block are shown in Figure 17. The OA, AMP2, CMP, and LOGIC blocks consume 360 nA, 440 nA, 100 nA, and 4.8 nA, respectively, and the SAM consumes a total of 1 μA of DC current. The clock generator (CLK_GEN) drew an average of 2.1 nA of current. The reduction in power consumption due to the SAM feature may vary according to the user scenario. The average current consumption is 14.3 μA when only the SF block without the SAM block is operated in 24-h normal power mode. When using the SAM function, in normal power mode, SF consumes 14.3 μA and SAM consumes 1 μA, so the total current is 15.3 μA. In low power mode, SF consumes 0.2 μA and SAM consumes 1 μA, so the total current is 1.2 μA. If there is a sound input for an average of 10% of the time of the day, it is used in the low power mode for 90% of the time and in the normal power mode for 10% of the time. Based on 10% of the current consumption of 15.3 μA in normal power mode and 90% of the current consumption of 1.2 μA in low power mode, the average current consumption is 2.46 μA. Without the SAM function, the current consumption is 14.3 μA, and with the SAM function, the average current consumption is 2.46 μA, so the current consumption can be reduced by 83%.
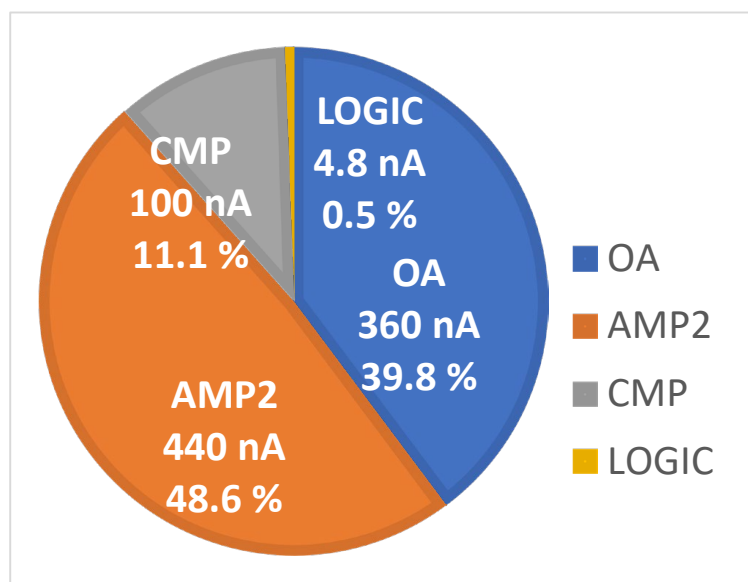

Figure 17. Breakdown of current consumption in SAM.

Table 1 summarizes the performance of the developed circuits and compares them with the results of the previous work. In 2020, Yang et al. announced a microphone with a Sound Activity Detector (SAD) function [13]. This work and Yang's previous work are designed using the same CMOS 0.18 μm process and have the same operating voltage of 1.8 V. Sound-activity monitoring features are designed to reduce microphone average power consumption in both this and previous works. As the architecture improves in this work, the current consumption is improved to 1 μA compared to the previous 2.5 μA. In the previous work, envelope detectors are used as a component of the SAM architecture [13]. According to the differential signal flow, two envelope detectors require components such as at least two amplifiers, two capacitors, two switches, and two leakage current elements. This work improves power consumption by eliminating envelope detectors and simplifying the SAM structure. In previous work, both inputs of SAM and SF were simultaneously connected to the MEMS port, which was able to reduce the sensitivity of MEMS by adding the input capacitance of SAM to the parasitic capacitance. To prevent this degradation in MEMS sensitivity, this work improves the architecture by connecting the output of SF to the input of SAM.

Table 1. Performance summary and comparison table with previous work.

Specification	This Work	TCAS-II'20 [13]
Technology	0.18 μm	0.18 μm
Supply voltage	1.8 V	1.8 V
Wake-up feature	YES	YES
Current consumption of SAM	1 μA	2.5 μA
SAM architecture	Amplifiers + Comparator	Amplifiers + Envelope Detector + Comparator
SAM input	SF output	MEMS device
Noise level at 1 kHz at low power mode	$72\ nV_{rms}/\sqrt{Hz}$	N.A.

4. Conclusions

A new technology was proposed to reduce the average power consumption of a microphone that is always operated. It was developed based on a scenario in which it operates only with minimal elements with low power when the sound is silent and then operates normally when the sound starts to be input. To operate such a scenario, a low-power sound-activity monitor (SAM) is required to generate a control signal that adjusts the power mode by detecting the presence or absence of sound. This work proposed a new SAM architecture with conventional envelope detectors removed, and the simplified SAM is designed based on the 0.18 μm CMOS process. Thanks to the simplified SAM architecture, SAM power consumption could be reduced to 1 μA. In an analog path consisting of SF, OA, and AMP2, low cutoff frequency can be adjusted between 30 Hz and 720 Hz, and high cutoff frequency has variable values from 1.4 kHz to 3.7 kHz. On the analog path, the amplification gain has a minimum of 55 dB (550 V/V) and a maximum of 63 dB (1395 V/V). In low power mode, the noise floor at 1 kHz is $72\ nV_{rms}/\sqrt{Hz}$, and the integrated noise between 600 Hz and 2 kHz has a value of $2.6\ \mu V_{rms}$. In addition, the input of SAM was connected to the output of SF to improve MEMS sensitivity, and the conversion delay from normal power mode to low power mode was added to improve the inefficiency caused by frequent conversion between power modes. The original intended function and performance were verified through simulation, and the proposed idea was confirmed to be feasible.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Meeker, M. Internet Trends. 2018. Available online: https://www.kleinerperkins.com/perspectives/internet-trends-report-2018 (accessed on 2 November 2022).
2. Five Voice Recognition Technology Trends & Applications. Available online: https://www.dolbeyspeech.com/blog/5-speech-voice-recognition-trends-applications (accessed on 2 November 2022).
3. Fortune Business Insights, Speech and Voice Recognition Market Projection. Available online: https://www.globenewswire.com/en/news-release/2022/05/31/2453438/0/en/Speech-and-Voice-Recognition-Market-US-28-3-Billion-by-2026-at-CAGR-of-19-8.html (accessed on 2 November 2022).

4. Yole Development 2017 Report, Acoustic MEMS and Audio Solutions. Available online: https://www.electronicspecifier.com/products/micros/audio-market-expected-to-be-worth-20bn-in-2022 (accessed on 2 November 2022).

5. MARKETANDMARKET, Microphone Market Global Forecast. Available online: https://www.marketsandmarkets.com/ResearchInsight/microphone-market.asp (accessed on 2 November 2022).

6. Du, D.; Odame, K. An adaptive microphone preamplifier for low power applications. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Seoul, Republic of Korea, 20–23 May 2012; pp. 660–663.

7. Chung, C.J.; Lu, C.; Rih, W.; Lee, C.; Shih, C.; Yeh, Y. An Ultra-low Power Voice Interface Design for MEMS Microphones Sensor. In Proceedings of the IEEE Sensors, Sydney, Australia, 31 October–3 November 2021; pp. 1–4.

8. Berti, C.; Malcovati, P.; Crespi, L.; Baschirotto, A. A 106 dB A-Weighted DR Low-Power Continuous-Time ΣΔ Modulator for MEMS Microphones. *IEEE J. Solid-State Circuits* **2016**, *51*, 1607–1618. [CrossRef]

9. Cho, S.; Kim, B.; Sim, J.; Park, H. Low-Power Small-Area Inverter-Based DSM for MEMS Microphone. *IEEE Trans. Circuits Syst. II Express Briefs.* **2020**, *67*, 2392–2396. [CrossRef]

10. Delbruck, T.; Koch, T.; Berner, R.; Hermansky, H. Fully integrated 500 μW speech detection wake-up circuit. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Paris, France, 30 May–2 June 2010; pp. 2015–2018.

11. Badami, K.M.H.; Lauwereins, S.; Meert, W.; Verhelst, M. A 90 nm CMOS 6μW power-proportional acoustic sensing frontend for voice activity detection. *IEEE J. Solid-State Circuits* **2016**, *51*, 291–302.

12. Cho, M.; Oh, S.; Shi, Z.; Lim, J.; Kim, Y.; Jeong, S.; Chen, Y.; Blaauw, D.; Kim, H.; Sylvester, D. A 142nW voice and acoustic activity detection chip for mm-scale sensor nodes using time-interleaved mixer-based frequency scanning. In Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, San Francisco, CA, USA, 11–15 February 2019; pp. 278–279.

13. Yang, Y.; Lee, B.; Cho, J.S.; Kim, S.; Lee, H. A Digital Capacitive MEMS Microphone for Speech Recognition with Fast Wake-Up Feature Using a Sound Activity Detector. *IEEE Trans. Circuits Syst. II Express Briefs.* **2020**, *67*, 1509–1513. [CrossRef]

14. Geronimo, G.; O'Connor, P.; Kandasamy, A. Analog CMOS peak detect and hold circuits. Part 1. Analysis of the classical configuration. *Nucl. Instrum. Methods Phys. Res. A* **2002**, *484*, 533–534. [CrossRef]

15. Kruiskamp, M.W.; Leenaerts, D.M. A CMOS peak detect sample and hold circuit. *IEEE Trans. Nucl. Sci.* **1994**, *41*, 295–298. [CrossRef]

16. Sharroush, S.M. Design of the CMOS inverter-based amplifier: A quantitative approach. *Int. J. Circuit Theory Appl.* **2019**, *47*, 1006–1036. [CrossRef]

17. Figueiredo, M.; Santos-Tavares, R.; Santin, E.; Ferreira, J.; Evans, G.; Goes, J. A Two-Stage Fully Differential Inverter-Based Self-Biased CMOS Amplifier with High Efficiency. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2011**, *58*, 1591–1603. [CrossRef]

18. Harrison, R.R.; Watkins, P.T.; Kier, R.J.; Lovejoy, R.O.; Black, D.J.; Greger, B.; Solzbacher, F. A Low-Power Integrated Circuit for a Wireless 100-Electrode Neural Recording System. *IEEE J. Solid-State Circuits* **2007**, *42*, 123–133. [CrossRef]

19. Zhang, F.; Holleman, J.; Otis, B.P. Design of Ultra-Low Power Biopotential Amplifiers for Biosignal Acquisition Applications. *IEEE Trans. Biomed. Circuits Syst.* **2012**, *6*, 344–355. [CrossRef] [PubMed]

20. Sun, Y.; Yu, X. Capacitive Biopotential Measurement for Electrophysiological Signal Acquisition: A Review. *IEEE Trans. Biomed. Circuits Syst.* **2016**, *16*, 2832–2853. [CrossRef]

21. Allen, P.E.; Holberg, D.R. Chapter 8 Comparators. In *CMOS Analog Circuit Design*, 3rd ed.; Oxford University Press: Oxford, UK, 2011.

22. Leakage Current-Based Delay Circuit. U.S. Patent US9667241B2, 30 May 2017.