Low Power Microphone Front-Ends

Lorenzo Crespi, Claudio De Berti, Brian Friend, Piero Malcovati, and Andrea Baschirotto

1 Introduction

After the invention of the first microphone in 1876, carbon microphones have been introduced in 1878 as key components of early telephone systems. In 1942, ribbon microphones were developed for radio broadcasting. The invention of the self-biased condenser or electret microphones (ECM) in 1962 represented the first significant breakthrough in this field. Indeed, electret microphones, ensuring high-sensitivity and wide bandwidth at low cost, have dominated the market for high-volume applications until the last decade, when MEMS microphones started to gain popularity [1].

The first microphone based on silicon micro-machining (MEMS microphone) was introduced in 1983. Thanks to the use of advanced fabrication technologies, MEMS microphones offer several advantages with respect to electret devices: better performance, smaller size, compatibility with high-temperature automated printed circuit board (PCB) mounting processes, and lower sensitivity to mechanical shocks. Moreover, MEMS microphones can be integrated together with the CMOS electronics on the same chip or, more commonly, within the same package [2], thus reducing area, complexity, and costs, while increasing efficiency, reliability, and

L. Crespi (✉) · C. De Berti · B. Friend
Synaptics, San Jose, CA, USA
e-mail: Lorenzo.Crespi@synaptics.com

P. Malcovati
University of Pavia, Pavia, Italy

A. Baschirotto
University of Milano-Bicocca, Milan, Italy

© Springer Nature Switzerland AG 2019

K. A. A. Makinwa et al. (eds.), *Low-Power Analog Techniques, Sensors for Mobile Devices, and Energy Efficient Amplifiers*,
https://doi.org/10.1007/978-3-319-97870-3_17

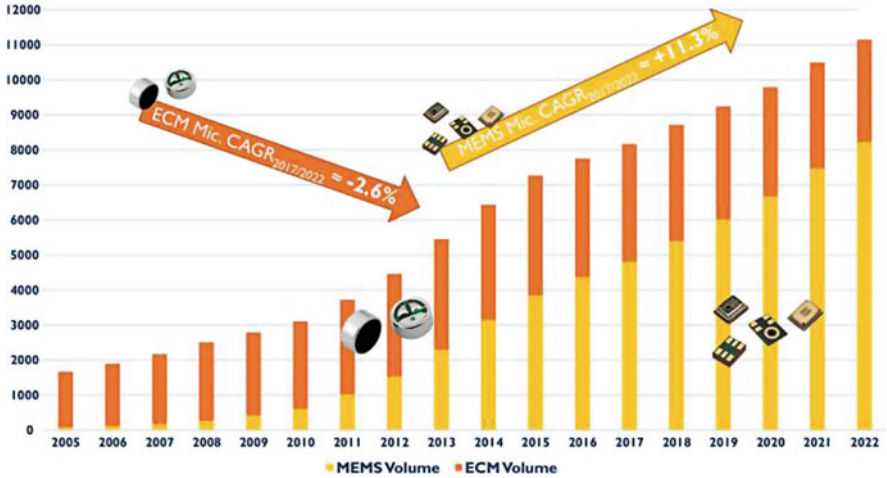


Fig. 1 The microphone market in million units since 2005. (Source: Acoustic MEMS and Audio Solutions 2017 Report, Yole Développement)

performance. As a result, around 2014, MEMS microphones surpassed ECMs in term of sold units, with an annual market size increase of more than 11%, as shown in Fig. 1.

MEMS microphones can be realized by exploiting different transduction principles, such as piezoelectric, piezoresistive, and optical detection. However, more than 80% of the MEMS microphones produced are based on capacitive transduction, since it achieves higher sensitivity, consumes lower power, and is more compatible with batch production.

The front-end circuit is of paramount importance for MEMS microphones, since it represents one of the most significant competitive advantages with respect to ECMs. Therefore, the development of high-performance front-end circuits has always progressed in parallel with the evolution of MEMS microphones [3–11]. This has led to a steady reduction of their power consumption, while maintaining or even improving their audio performance, such as signal-to-noise ratio (SNR), dynamic range (DR), and total harmonic distortion (THD). This trend is mainly driven by portable applications, whose audio-related functionality has expanded significantly. For example, voice interfaces are becoming pervasive. A growing number of people now talk to their mobile devices, asking them to send e-mails and text messages, to search for directions, or to find information on the internet. These functions require continuous listening, thus introducing severe constraints on the power consumption of the microphone modules. Low power consumption is, therefore, the key design goal of modern front-end circuits for MEMS microphones.

2 Capacitive Microphones

A microphone is a sensor that translates a perturbation of air pressure, i.e., sound, into an electrical quantity. In a capacitive microphone, pressure variations cause the vibration of a mechanical mass, which is transformed into a capacitance variation. Sound pressure is typically expressed in dB_{SPL} (sound-pressure-level).

A sound pressure of 20 μPa, corresponding to 0 dB_{SPL}, is generally accepted as the auditory threshold (the lowest amplitude of a 1-kHz signal that a human ear can detect). The sound pressure levels of a face-to-face conversation range between 60 dB_{SPL} and 70 dB_{SPL}. This rises to 94 dB_{SPL} if the speaker is at a distance of 1 inch from the listener (or the microphone), which is the case, for example, in mobile phones. Therefore, a sound pressure level of 94 dB_{SPL}, which corresponds to 1 Pa, is used as a reference for acoustic applications. The performance parameters for acoustic systems, such as the SNR, are typically specified at 1-Pa and 1-kHz. Some additional examples of typical SPL levels are shown in Fig. 2.

2.1 MEMS Microphones

A MEMS capacitive microphone, whose simplified structure is shown in Fig. 3, basically consists of two conductive plates at a distance *x*. The top plate, in this

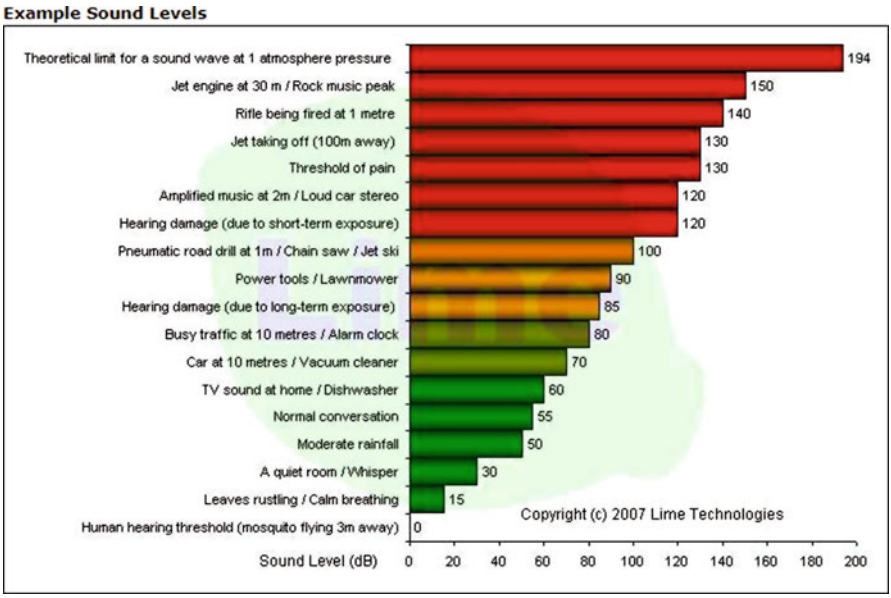


Fig. 2 Example sound levels in dB_{SPL}

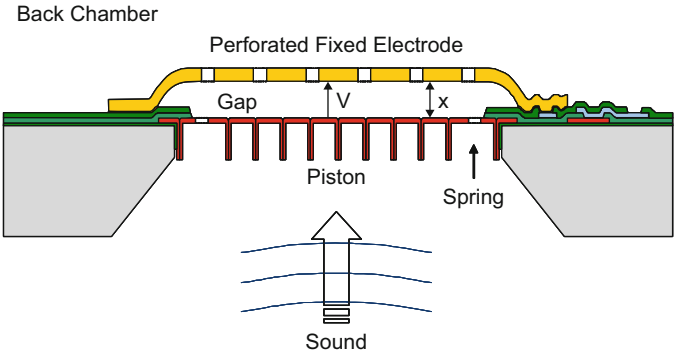


Fig. 3 Basic structure and working principle of a MEMS capacitive microphone

case, is fixed and cannot move, while the bottom plate is able to move in response to sound pressure, producing a variation of x (Δx) with respect to its steady-state value (x_0), proportional to the instantaneous pressure level (P_S). Different arrangements of the electrodes and fabrication solutions are possible, but the basic principle does not change [12–18].

The capacitance of a MEMS microphone is then given by

$$C\left(P_S\right) = \frac{\varepsilon_0 A}{x\left(P_S\right)} = \frac{\varepsilon_0 A}{x_0 + \Delta x\left(P_S\right)} \tag{1}$$

where A is the area of the smallest capacitor plate and ε_0 is the vacuum dielectric permittivity.

The MEMS microphone capacitor is initially charged to a fixed voltage V_B, with a charge $Q = C_0 V_B$, where C_0 is the capacitance value in the absence of sound ($x = x_0$). Therefore, assuming a linear relation between the sound pressure variation ΔP_S and the displacement Δx ($\Delta x = -k\Delta P_S$), the capacitance variation leads to a voltage signal (ΔV) across the microphone, given by

$$\Delta V = \frac{Q}{C(P_S)} - \frac{Q}{C_0} = \frac{Q\Delta x}{\varepsilon_0 A} = -\frac{kC_0 V_B}{\varepsilon_0 A}\Delta P_S = -\kappa \Delta P_S \tag{2}$$

where κ denotes the sensitivity of the microphone. In order to avoid degradation of the voltage signal ΔV, the input impedance of the front-end circuit must be extremely large, thus ensuring that Q remains constant.

In practical implementations, a MEMS microphone is not just a capacitor—some additional parasitic components also have to be taken into account. The equivalent circuit of an actual MEMS microphone is shown in Fig. 4.

Besides the variable capacitance $C(P_S)$, the equivalent circuit includes two parasitic capacitances, C_{P1} and C_{P2}, connected between each plate of the MEMS microphone and the substrate, as well as a parasitic resistance R_P, connected in parallel to $C(P_S)$. The value of these parasitic components depends on the specific

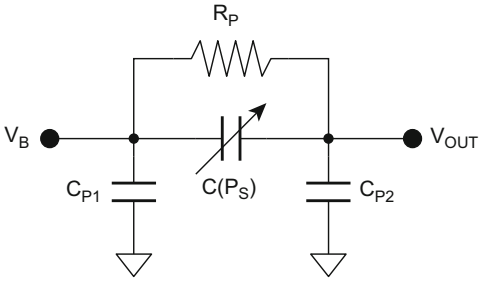


Fig. 4 Equivalent circuit of an actual MEMS capacitive microphone

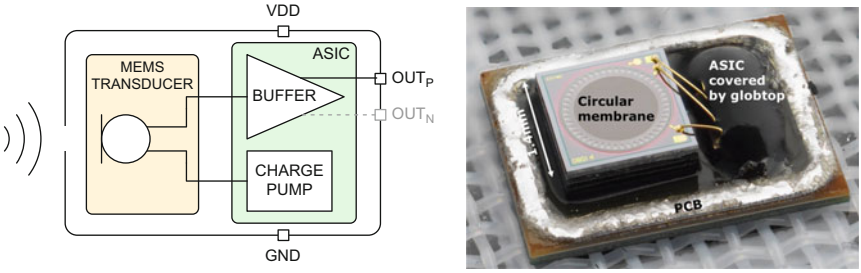


Fig. 5 Typical commercial MEMS microphone module

implementation of the microphone, but typically C_{P1} and C_{P2} are in the order of a few pF, while R_P is in the GΩ range.

2.2 MEMS Microphone Modules

The extremely large source impedance of a capacitive MEMS sensor makes its output signal very susceptible to EM interference and attenuation by routing parasitics. In most systems, it would thus be impractical to route the unbuffered MEMS sensor output, via wires or PCB traces, to the System-on-Chip (SoC) responsible for digitizing and processing it.

A MEMS microphone sensor is typically co-packaged with a small ASIC including biasing and buffering circuits, as shown in Fig. 5. A charge-pump-up converts the supply voltage V_{DD} to generate the MEMS bias voltage V_B. Since the sensor sensitivity is proportional to its bias voltage as shown in (2), V_B is set to a relatively high voltage, typically in the 8–12 V range. V_B is limited on the high side to a critical voltage called the pull-in voltage, at which the MEMS membrane collapses and the device ceases to operate properly.

A simple low-noise amplifier with a very high input impedance then generates a buffered version of the microphone signal, which can be routed via wires or PCB traces to the processing SoC. In its simplest form, this amplifier could be implemented by using a single FET transistor. The output of the microphone module

is typically single-ended, but balanced differential outputs are becoming more commonly available, in consideration of the higher performance with negligible additional power consumption.

2.3 Performance of MEMS Microphone Modules

Performance of commercial microphone modules is generally specified by the following key parameters:

Sensitivity The rms voltage produced at the microphone output in response to a 94-dB_{SPL}, 1-kHz sinusoidal input, expressed in dB_V. For modern MEMS sensors, microphone sensitivity typically ranges from -32 dB_V to -42 dB_V.

Sensitivity Tolerance This is a particularly critical parameter for microphone arrays, where mismatched gains can degrade performance of beam-forming and other voice processing algorithms. State-of-the-art MEMS microphones typically achieve $\pm 1\%$ sensitivity matching. This is a significant improvement over ECM microphones that are usually rated at $\pm 3\%$.

Signal-to-Noise Ratio (SNR) The ratio between the output produced by a reference 1-kHz signal at 94 dB_{SPL} and the residual output noise floor with no input, integrated over 20 Hz–20 kHz band with A-weighting. Many recent MEMS microphones achieve SNRs in the 60–70 dB range, with best-in-class modules now approaching SNRs of 75 dB. The best ECM microphones still hold a slight advantage over MEMS devices in this category, reaching up to 80 dB SNR at the expense of much larger physical dimensions.

Acoustic Overload Point (AOP) The sound pressure level at which microphone THD equals 10%. It indicates the maximum acoustic level that the microphone can process without drastically distorting the signal. Typical AOP levels for current MEMS microphones are 120–130 dB_{SPL}, with some microphones now achieving 135–140 dB_{SPL}. The trend in recent years has been toward rapidly increasing AOPs. While the benefit of reaching AOPs larger than the human threshold of pain (see Fig. 2) may seem questionable, at least in the context of consumer electronic products, a high AOP is actually very useful to prevent microphone saturation from wind noise, proximity to a powerful loudspeaker, or from low-frequency thump-like signals, which can occur in a car interior during door closing, or while a train is going through a tunnel, and so on. A temporary microphone saturation can be disruptive to adaptive voice-processing algorithms, such as the ones used in acoustic noise cancelling (ANC) headphones, and should be avoided.

Distortion (THD or THDN) Typically measured at 1 kHz, and at different sound pressure levels, depending on the manufacturer, THD typically ranges from 1% to 0.04%.

Output Impedance Typically, in the 50–1000 Ω range.

Power Supply Rejection (PSRR or PSR) Both indicate the capability of the ASIC to reject spurious noise on the supply voltage; the main difference is that the PSRR is expressed as a dB ratio, while the PSR is expressed in dB_V or dB_V A-weighted ($dB_{V\text{-}Aw}$). Test conditions vary among manufacturers, but generally a 217 Hz or 1 kHz, 100-mV_{pp} square wave or sine wave is injected as supply noise. The typical range for PSRR is 45–75 dB.

3 Microphone Front-End Architecture and Specifications

The interface circuit for a MEMS module reads out an analog signal and converts it to the digital domain. The system diagram for a typical front-end circuit for a MEMS capacitive microphone module is shown in Fig. 6, for both single-ended and differential microphones. The circuit consists of a programmable-gain preamplifier (PGA) followed by an analog-to-digital converter (ADC). The input of the preamplifier is typically AC-coupled to remove the DC voltage at the microphone output. The RC network created by the AC coupling can also be useful as a high-pass filter (HPF) to filter out low-frequency noise, such as those generated by wind and other undesirable acoustic sources.

In the case of a single-ended microphone output, it is best to AC couple the ground terminal of the microphone to the negative input of the PGA, in order to reject common-mode interference that may couple into the wiring or the PCB traces. A series resistor on the ground line is often used to equalize the impedance level on the negative line, which improves RFI rejection [19]. Series ferrite beads

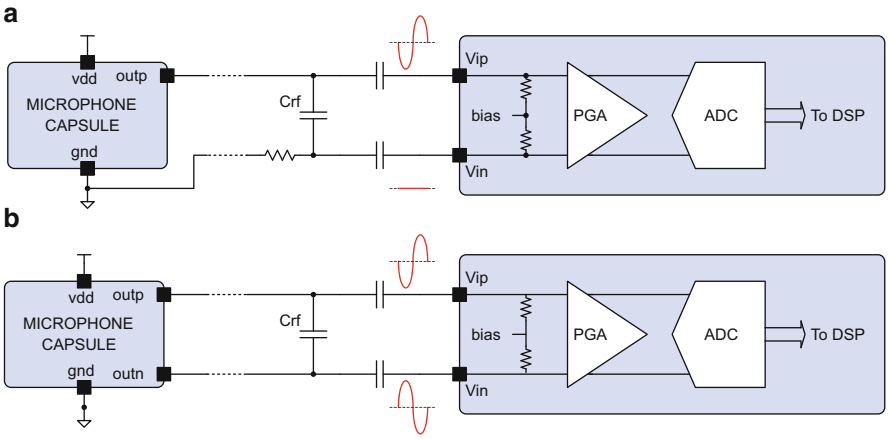


Fig. 6 Typical block diagram of the front-end circuit for a MEMS microphone module: (**a**) single-ended microphone; (**b**) differential microphone

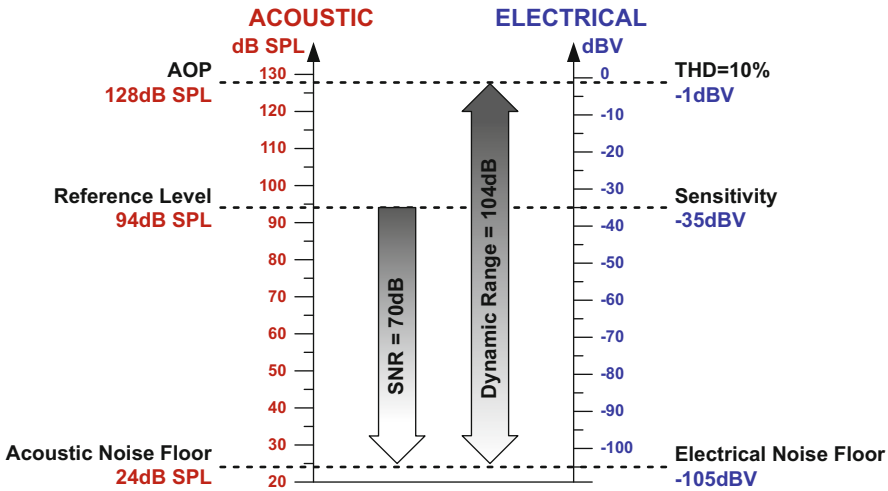


Fig. 7 Microphone parameters in acoustic vs. electrical domain

and/or a small RF shunt capacitor are also commonly placed to reduce RF noise in traces [20].

3.1 Interface Requirements

In general terms, the fundamental requirement of a microphone interface is to digitize the analog signal from the microphone *without significantly degrading its quality*. Since the microphone module is usually selected by the system manufacturer based on various criteria (cost, performance, physical dimensions, manufacturability, business relationships, etc.), it is imperative for a general-purpose microphone interface to be able to efficiently couple with a wide range of state-of-the-art commercial microphones modules. The following section describes how the key microphone parameters can be translated into electrical specifications for its interface circuits. The performance quality of different microphone front-ends is of course in trade-off with the power consumption. Generally, the higher the power consumption, the better is the performance.

Acoustic to Electrical Domain
Figure 7 illustrates the relationship between microphone sensitivity, SNR, AOP, and DR, in both acoustic and electrical domains, for a hypothetical microphone with −35-dB sensitivity, 70-dB SNR, and 128-dB_{SPL} AOP.

In Fig. 8, the SNR/sensitivity/AOP specifications for available MEMS modules from various manufacturers are collected and translated into noise floor and voltage

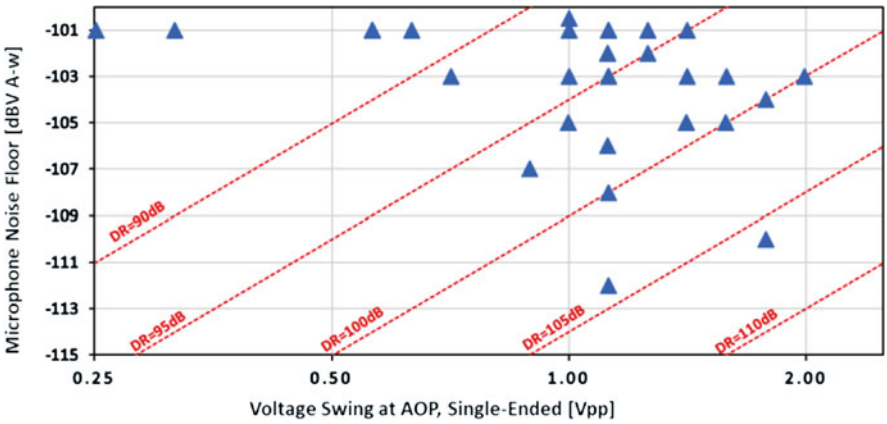


Fig. 8 Voltage swing and electrical noise floor for commercial MEMS microphones

swing at the microphone output. The voltage swing is shown as peak-to-peak single-ended, as this is the most useful information to determine headroom requirements for the preamplifier. From this chart, a few key parameters for the interface circuit can be extracted:

Max Input Voltage Swing While conventional ECM (and earlier MEMS) microphones typically produce a signal in the order of 100 mV_{pp} or less, recent MEMS microphones with high AOP and sensitivity can generate a significantly larger signal, in the order of 1–2 V_{pp} single-ended or 2–4 V_{pp} differential. A general-purpose microphone interface should be able to handle such signal without distorting; depending on the circuit architecture, this can entail using a higher supply voltage for the input stage of the preamplifier relative to the rest of the interface circuitry.

Input-Referred Noise and Dynamic Range Many high-end MEMS microphones have an output noise floor close to -105 $dB_{V\text{-}Aw}$, with the best in class reaching up to -112 $dB_{V\text{-}Aw}$. Therefore, a high-performance microphone interface should have an input-referred noise lower than -118 $dB_{V\text{-}Aw}$, in order to avoid degradation of the overall SNR and DR (this, of course, requires higher power).

Preamplifier Gain The preamplifier buffers the signal from the microphone and scales its amplitude to match the full-scale of the ADC. In principle, a fixed preamplifier gain is sufficient; however, meeting all worst-case requirements for

voltage-swing and input-noise simultaneously is a very challenging proposition. Handling 2 V_{pp} full-scale with a –118 $dB_{V\text{-}Aw}$ noise floor requires an ADC dynamic-range of 115 dB, which can be expensive in terms of die area and power. To alleviate the ADC requirements, a preamplifier with variable gain is generally employed to compensate for different microphone sensitivities. The low-end of the preamp gain range is determined by the largest microphone signals, as discussed in the previous paragraph. Assuming an ADC full-scale of 1 V_{rms} differential, and a max input swing of 2 V_{pp} single-ended, a minimum preamplifier gain of 3 dB is adequate. At the high-end, preamplifiers have traditionally implemented gains in the 20–40 dB range; however, given the recent increase in microphone AOP levels, this is no longer possible. As shown in Fig. 8, most modern MEMS microphones can generate at least 0.5–1 V_{pp} near AOP, which limits the max usable gain to 12–15 dB. The preamplifier gain steps should be 3 dB or less to allow tailoring the interface characteristics to the specific microphone used in the system.

AC Versus DC Coupling AC coupling is prevalent because it blocks the unknown DC voltage across the microphone with no power consumption or performance impact. This is typically implemented with an external and expensive capacitor in the order of a few μF to keep the high-pass pole in the order of 1 Hz. DC coupling is recently being introduced for applications that have stringent constraints for PCB area or BOM cost. A few solutions have been proposed to implement DC coupling [21–24]. However, a trade-off between power consumption, SNR performance, and/or die area is generally unavoidable when designing DC-coupled preamplifiers. This chapter focuses on AC-coupled interfaces.

Input Impedance The source impedance of MEMS microphones typically ranges from 200 Ω to 1 kΩ (or 2.2 kΩ if ECM mics are included). Even MEMS microphones with low-output impedance are often current-limited and unable to drive their peak signal into heavy resistive loads. To avoid significant attenuation and distortion of the microphone signal, a general-purpose preamplifier must present an input impedance in the order of 10 kΩ or larger. The presence of an AC-coupling capacitor on the microphone inputs adds further restrictions to the preamplifier input impedance, due to the HPF formed with the input resistance of the stage.[1]

Linearity Given that most microphones are limited to ≥0.04% THD (−68 dB), the linearity requirement for the interface circuit is fairly relaxed compared to other parameters. A THD < −75 dB is typically sufficient for most applications.

[1]To achieve 20 Hz cutoff frequency with a 10-kΩ input resistance, the AC-coupling caps must be of the order of 1 μF. While 1-μF ceramic capacitors are widely available even in very small form factor, their large voltage coefficient can create a significant nonlinearity at low frequencies. For this reason, it is strongly preferable to utilize capacitors in the order of 10 nF, which requires a preamplifier input impedance in the order of 1 MΩ.

In the following, the circuit and system solutions for each block (PGA and ADC) will be introduced, emphasizing the trade-off between power consumption and performance.

4 Preamplifier Design

A conventional preamplifier consists of a resistive feedback operational amplifier with a large input resistor, as shown in Fig. 9.

This architecture is used in many commercial products, since it is quite simple and ensures good linearity even with large input signals, but it has several limitations to realize wide gain range. Indeed, to avoid attenuating the microphone signal, the input resistor should be large, thus requiring an even larger feedback resistor. As a result, both the preamplifier area and input referred noise become excessive. To overcome these limitations, a convenient solution is to use a preamplifier based on a transconductance input stage, as shown in Fig. 10, thus achieving both high-input impedance and high gain range without requiring large resistors that contribute noise.

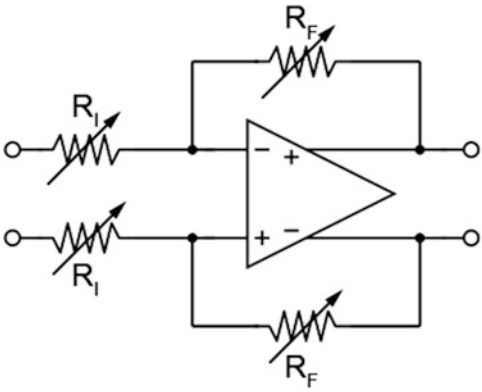


Fig. 9 MEMS microphone preamplifier with resistive feedback

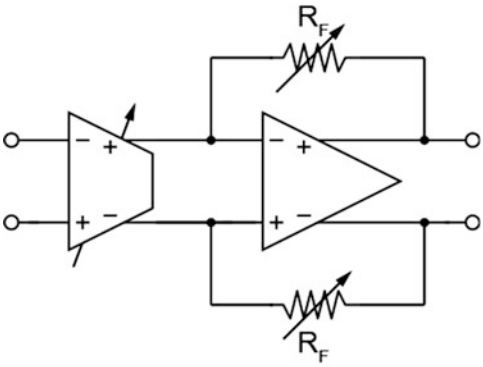


Fig. 10 MEMS microphone preamplifier with transconductance input stage

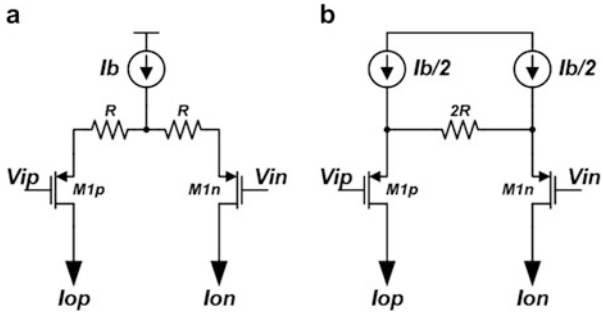


Fig. 11 Simple implementation of transconductor with lumped tail current (**a**) and with split tail current (**b**)

4.1 State of the Art: Transconductance Amplifier

The efficiency of the circuit of Fig. 10 depends on the implementation of the input transconductance stage, which must combine low power consumption with wide gain programmability.

In its simplest form, a linearized transconductor can be implemented as a source-degenerated differential pair biased at a constant current I_b. Its total transconductance is $G_m = g_m/(1 + g_m R)$, which can be approximated to R when $R \gg 1/g_m$.

Figure 11 shows two implementations of a source-degenerated differential pair. The two solutions provide the same input/output transfer function, but version (b) is often preferred because of its improved voltage headroom, given the fact that current I_b does not flow through the degeneration resistors. However, version (a) presents a fundamental advantage noise-wise: the noise current associated with the bias current I_b splits equally between I_{op} and I_{on} when $V_{ip} \approx V_{in}$ (small signal conditions) and becomes a common-mode noise component that is rejected by the following trans-resistance stage. On the other hand, in version (b), the two tail currents produce uncorrelated noise currents which are added to the differential signal current. Moreover, their mismatch would produce offset. This makes structure (a) the better choice for audio preamplifiers.

At full-scale signal conditions, the two circuits are almost equivalent as the noise from I_b is steered completely into I_{op} and I_{on} and is added to the differential signal current. With a full-scale sinusoidal input, version (a) retains a 3-dB advantage over version (b).

To further enhance the transconductor linearity, transistors M_{1p} and M_{1n} can be supplemented with feedback structures that decrease their output resistance and generate a more accurate copy of $V_{ip} - V_{in}$ voltage across resistor R_1. A well-known example based on the super-source-follower (SSF) is shown in Fig. 12b. This simple circuit is very effective in this application, and relative to (a), it biases

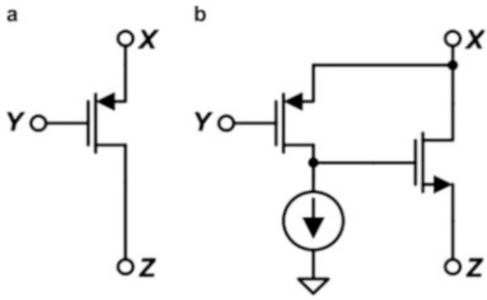


Fig. 12 From single transistor to Type-II current conveyor

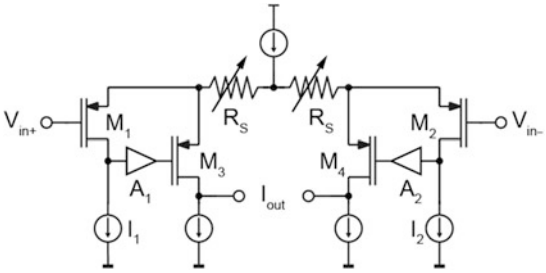


Fig. 13 Schematic of the transconductance stage for a MEMS microphone preamplifier

the input transistor M_1 at constant current, thus maintaining a signal-independent V_{gs} and dividing the impedance on node X by a factor $g_m r_0$.

A good example of a MEMS microphone preamplifier based on this technique has been proposed in [25] and is shown in Fig. 13. Transistors M_1 and M_2, current sources I_1 and I_2, and inverting amplifiers A_1 and A_2 form an active feedback loop for improving linearity. The effective transconductance of the stage is determined by the source degeneration resistances R_S ($g_m = 1/R_S$). Compared to a conventional degenerated differential pair, the linearity and gain accuracy of this transconductor are enhanced by an additional factor $g_{m1,2}A_{1,2}R_{X,Y}$, where $g_{m1,2}$ is the transconductance of M_1 and M_2, $A_{1,2}$ is the gain of the inverting amplifiers, and $R_{X,Y}$ is the impedance at node X or Y. With these additional design parameters, the input-referred noise, the linearity, and the gain accuracy can be optimized independently. The noise effect of M_3, M_4, and R_S is the same as in a conventional degenerated differential pair, but the high-loop gain of the active feedback loop helps to reduce the input-referred noise of all the components except transistor M_1, M_2, I_1, and I_2. Compared to a conventional transconductor, this circuit achieves better linearity and gain accuracy with equal or lower power consumption.

The THD + N of a preamplifier based on the scheme shown in Fig. 13, featuring a gain range from 22 dB to 42 dB is illustrated in Fig. 14. This preamplifier consumes 350 μW.

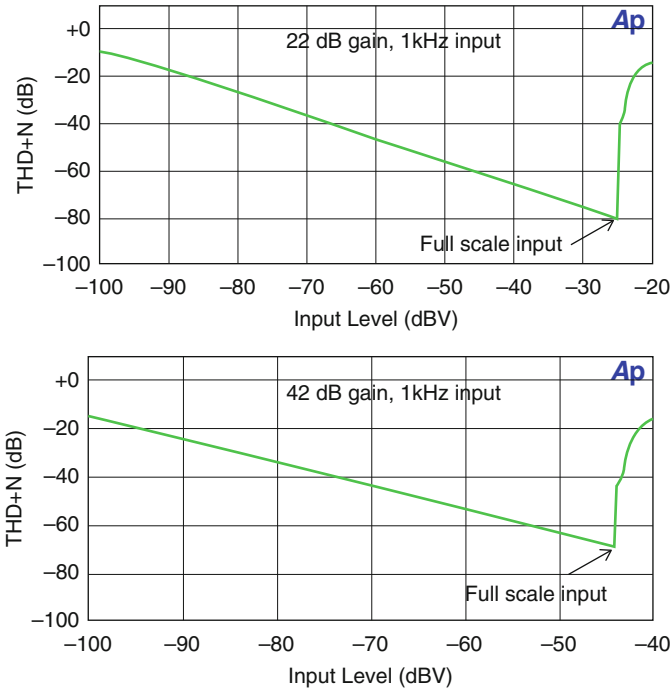


Fig. 14 Measured THD + N of a MEMS microphone preamplifier with transconductance input stage

4.2 Improving the Transconductance Amplifier

"Class-H" Adaptive Biasing

Further improvement in terms of efficiency can be achieved with adaptive biasing techniques, which allow the average power consumption of audio circuits to be reduced and takes advantage of the bursty nature of voice/audio signals. Some authors have proposed a bandwidth-adaptive preamplifier [26]. Examples of amplitude-adaptive amplifiers have been proposed in [27].

A conventional source-degenerated transconductor is biased in Class-A, with a constant current equal to or larger than the peak output current. However, when the incoming signal has small amplitude, the biasing current can be temporarily reduced without incurring any performance penalties. The amount of instantaneous bias current is controlled by an envelope detector circuit which tracks the amplitude of the input signal. This principle can be seen as the current-domain analog of traditional class-H voltage amplifiers. An envelope detector that can be used to adjust the tail current of the main transconductor is shown in Fig. 15.

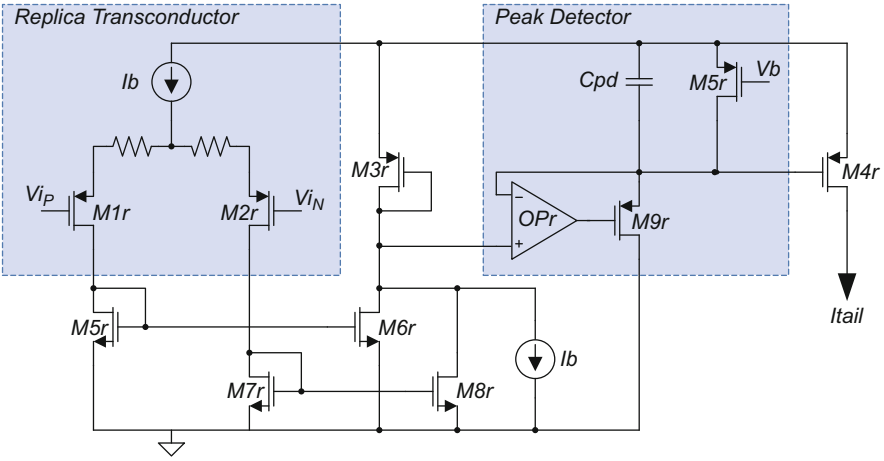


Fig. 15 Envelope detector used to generate the transconductor bias current

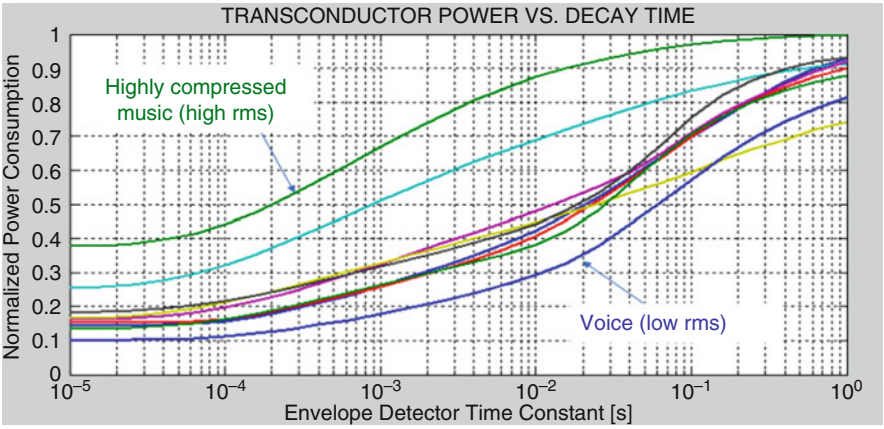


Fig. 16 Transconductor power vs. envelope detector time constant

A scaled version of the main transconductor generates a differential current proportional to the input signal, which is then rectified, converted to voltage-mode by transistor M_{3r}, and processed by a peak detector with a long decay time-constant in the millisecond range. The leaky element of the peak detector is implemented by a long-channel p-channel transistor M_5 biased in deep sub-threshold region. The output of the peak detector is then converted back to current I_{tail} by transistor M_{4r}.

A long time constant in the peak detector is useful to filter audio-band components from current I_{tail}, which could degrade overall THD due to the finite CMRR of the main transconductor. However, a trade-off exists between THD and power efficiency: a longer time constant keeps the PGA operating at high bias currents for a larger percentage of time. Figure 16 shows the theoretical transconductor power

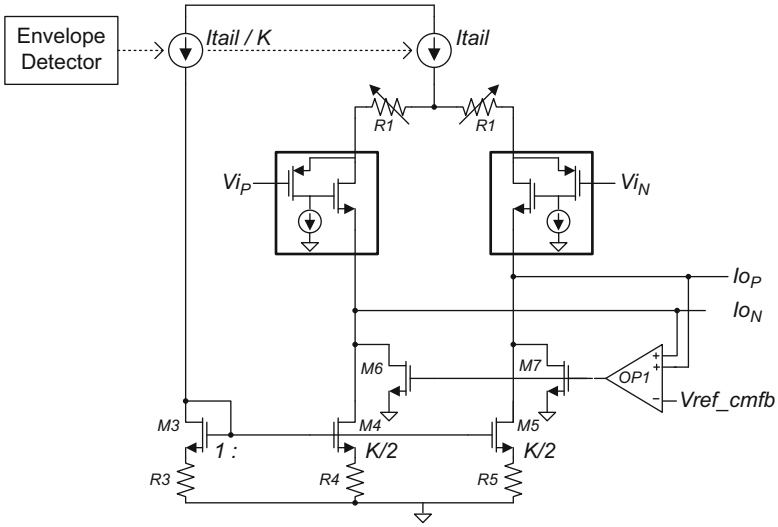


Fig. 17 Top level of transconductance amplifier

consumption vs. time constant, for various speech and music signals, normalized to the power consumption of an ideal Class-A transconductor. For a 10-ms time constant, the power savings from Class-H operation range from 12% (green curve, highly compressed music) to 71% (blue curve, speech).

Main Transconductor Circuit

The overall circuit for the transconductor is shown in Fig. 17. The variable tail current from the envelope detector is mirrored by source-degenerated n-channel transistors M_3, M_4, and M_5 to remove the common-mode component of I_{tail} from the output currents. Since the mirroring operation unavoidably introduces errors, a residual common-mode current exists and is cancelled by the common-mode feedback loop formed by OP_1, M_6, M_7.

Transconductor Gain-Selection Switches

The PGA gain is selected by switching the amount of degeneration resistance R_1. This optimizes noise vs. signal amplitude and, hence, maximizes efficiency. The switched resistor array is shown in Fig. 18. Since the switches are in series with the poly resistors and carry signal-dependent current, the linearity of the switch resistance directly impacts the THD performance of the PGA.

The voltage on the switch source V_{tail} is a rectified and level-shifted version of the input signal, which makes it impractical to implement the ON switches with p-channel transistors biased at $V_{\text{g}} = 0$, unless an extremely large W/L is chosen. Instead, the gate of the ON switches is biased at voltage $V_{\text{bON}} = V_{\text{tail}} - R_{\text{LS}} I_{\text{bLS}}$, therefore achieving a constant-V_{gs} biasing that makes the switch resistance nearly constant across signal swing. Current I_{bLS} is chosen to be $\ll I_{\text{b}}$.

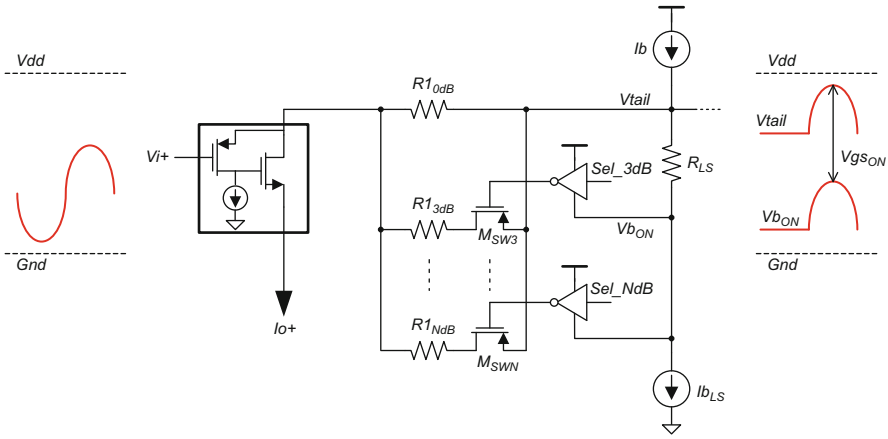


Fig. 18 Biasing of gain-selection switches

Supply Voltage Selection

Power consumption in the PGA can be minimized by selecting the most appropriate supply voltage for a given PGA gain setting. In most battery-powered systems, at least two power supplies are available: the battery itself (with a typical value of 3.7 V for Li-ion batteries) and one or more regulated supplies whose voltage depends on technology selection.

Low PGA gain is used for highly sensitive microphones that can output as much as 2 V_{pp} single-ended. In this case, the battery voltage should be used to maximize headroom. One problem with this approach is that the battery voltage is variable and generally quite noisy, due to its connections to DC/DC converters, RF power amplifiers, etc. Unless the tail current of the transconductor is designed to achieve very high PSRR, it is advisable to insert an LDO between the battery and the PGA supply.

Only the transconductor stage needs the higher supply voltage; the trans-resistance stage that follows can always be operated at the lower supply voltage.

For gains of 12 dB (signal ≤ 0.25 V_{rms}) or more, the signal swing is low enough to allow operation of the transconductor at 1.8 V.

The DC bias voltage at the transconductor input must be adjusted with the supply voltage, in order to keep the signal swing centered in the linear region of the transconductor.

Current Sources with Variable Source-Degeneration Resistors

A trade-off between noise and headroom exists when sizing the source degeneration resistors used for the noise-sensitive current sources: for a given current level, higher degeneration resistance means lower $1/f$ noise and higher voltage headroom.

When the PGA operates in its lowest gain setting (high-sensitivity microphone), the large signal swing requires using a minimal amount of resistive degeneration. This is acceptable since the input-referred noise can also be increased in large

signal conditions. As the gain increases, the headroom requirements become more relaxed, while the noise requirements become more stringent, and it is appropriate to progressively increase the amount of source degeneration resistance.

5 A/D Converter

The ADC in MEMS microphone front-end circuits is typically implemented with a $\Sigma\Delta$ Modulator ($\Sigma\Delta$M), which exploits oversampling to achieve the required DR. In particular, continuous-time (CT) $\Sigma\Delta$Ms represent the most promising solution for minimizing power consumption, since they require operational amplifiers (op-amps) with lower bandwidth with respect to switched-capacitor (SC) $\Sigma\Delta$Ms, which have been traditionally used. The Schreier figure of merit, defined as $\mathrm{FoM_S} = \mathrm{DR} + 10\log(B/P)$, B being the bandwidth and P the power consumption, is a useful indicator to compare different ADC solutions. Figure 19 shows the values of $\mathrm{FoM_S}$ of recently published ADCs as a function of the Nyquist frequency, $F_\mathrm{N} = 2B$.

5.1 *State of the Art: Continuous-Time $\Sigma\Delta$ Modulator*

In the audio field ($B = 20$ kHz), best-in-class performance ($\mathrm{FoM_S} = 180$ dB) has been achieved with the third-order CT $\Sigma\Delta$M with 15-level quantizer, whose block diagram is illustrated in Fig. 20. It achieves excellent efficiency thanks to several circuit and system choices as follows [28].

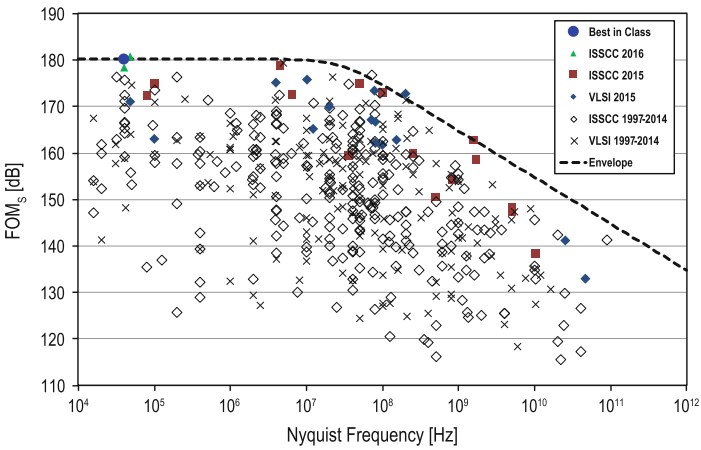


Fig. 19 ADC state of the art based on $\mathrm{FoM_S}$ from [29]

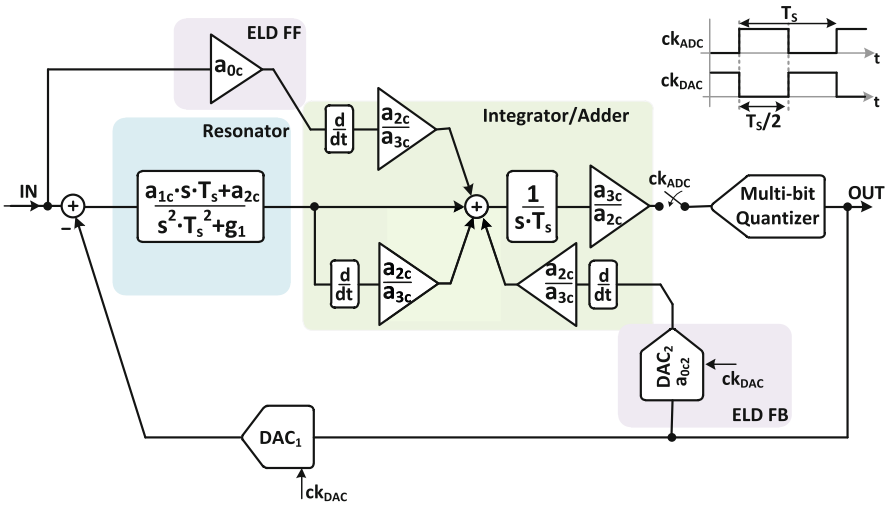


Fig. 20 Block diagram of the CT ΣΔM

The loop filter of the CT ΣΔM consists of a resonator (second-order transfer function) followed by an integrator. A local feedback DAC around the quantizer (DAC_2) and a dedicated feedforward path are used for compensating the excess loop delay (ELD). The feedforward paths of the loop filter and the local ELD feedback are differentiated and added at the input of the integrator, in order to avoid an active adder at the input of the quantizer. The multi-bit quantizer drives a 15-level DAC (DAC_1) with dynamic element matching (DEM) to close the main feedback loop of the CT ΣΔM.

The schematic of the active-RC implementation of the CT ΣΔM is shown in Fig. 21. The resonator is implemented using a single op-amp, and no active adder is used at the input of the quantizer, thus, requiring only two op-amps for implementing the third-order loop-filter transfer function. The local feedback DAC for ELD compensation is implemented with a SC structure, whereas the main feedback DAC is realized with a three-level (−1, 0, 1) current-steering topology, which guarantees minimum noise for small input signals. Indeed, with the three-level topology, the unused DAC current sources are not connected to the resonator input and, hence, they do not contribute to the CT ΣΔM noise. The multi-bit quantizer is realized with 14 identical differential comparators and a resistive divider from the analog power supply for generating the threshold voltages. The values of the passive components used for implementing the CT ΣΔM are summarized in Table 1. The value of R_i has been chosen as low as 47 kΩ to fulfill the thermal noise requirements, while R_1, R_3, R_4, C_1, C_2, C_f, and C_4 are obtained consequently to achieve the desired CT ΣΔM coefficients. Eventually, resistor R_i can be removed if the preamplifier is realized with a transconductor which provides directly an output current. Both op-amps are realized with a two-stage, Miller compensated topology in which transistor size and

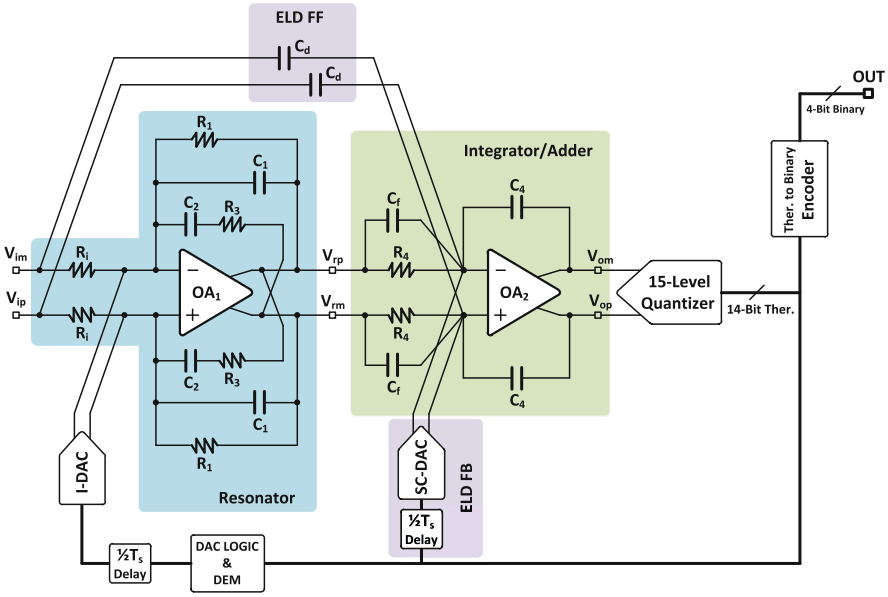


Fig. 21 Schematic of the active-RC implementation of the CT ΣΔM

Table 1 Values of the passive components used for implementing the CT ΣΔM

Resistor	Value	Capacitor	Value
R_i	47 kΩ	C_1	18.5 pF
R_1	5.7 MΩ	C_2	18.7 pF
R_3	57 kΩ	C_f	2.1 pF
R_4	1 MΩ	C_4	1 pF

bias current are sized to fulfill the noise requirements (the values in the second op-amp are scaled with respect to the first one, since its noise contribution is negligible).

The CT ΣΔM has been fabricated using a 0.16-μm CMOS technology. The micrograph of the 0.21-mm^2 chip is illustrated in Fig. 22. Figure 23 shows the measured SNDR as a function of the input sinusoidal signal amplitude at 1∼kHz. The full-scale input signal (0 dB_{FS}) corresponds to 1 V_{rms} differential. The achieved DR is 106 dB (A-weighted), corresponding to an ENOB of about 17 bits, whereas the peak SNDR is 91.3 dB. The change of slope in the SNDR curve for input signal amplitudes larger than −17 dB_{FS} is due to the increased current-steering DAC noise when more than 1 three-level DAC element is used (acceptable for the microphone application, where the performance for large input signals is limited by the microphone itself).

The CT ΣΔM output spectra obtained with −60 dB_{FS} and − 1 dB_{FS}, 1-kHz input signals are shown in Fig. 24. As expected, at −1 dB_{FS}, the noise floor increases of about 10 dB with respect to −60 dB_{FS}, due to the increased DAC noise. Figure 25 shows the measured inherent anti-aliasing properties of the CT ΣΔM. The spectral

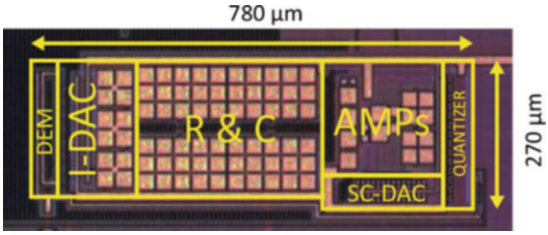


Fig. 22 Chip micrograph of the CT ΣΔM

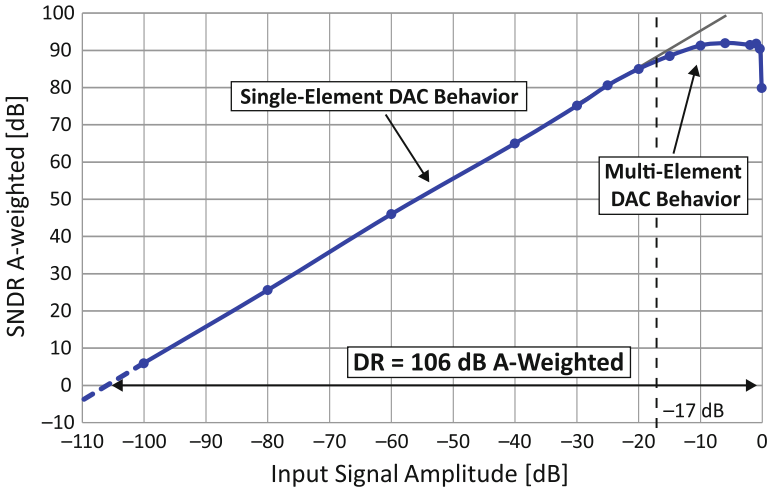


Fig. 23 Measured SNDR of the CT ΣΔM vs. input signal amplitude

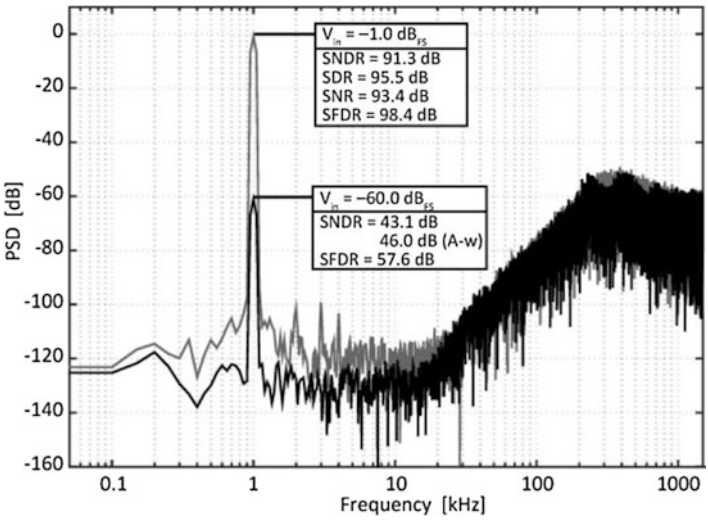


Fig. 24 Measured output spectra of the CT ΣΔM with −60 dB_{FS} and − 1 dB_{FS}, 1-kHz input signals

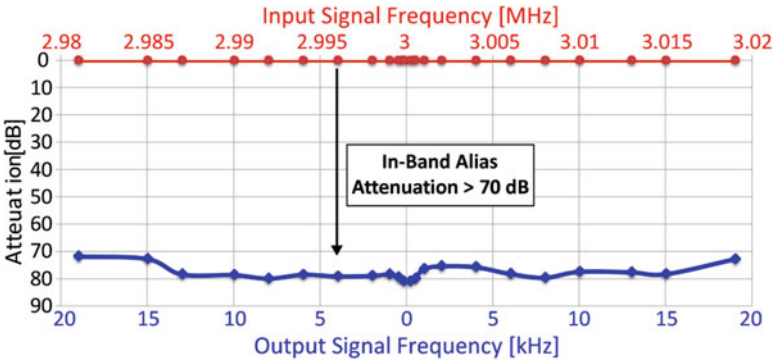


Fig. 25 Measured anti-aliasing properties of the CT ΣΔM

Table 2 Performance summary of the CT ΣΔM

Parameter	Value
Technology [nm]	160
Architecture	CT
Supply voltage [V]	1.6
Power consumption (*P*) [mW]	0.39
Bandwidth (*B*) [kHz]	20
Oversampling ratio (OSR)	75
Area [mm^2]	0.21
Peak SNDR [dB]	91.3
Dynamic range (DR) [dB]	103.1
Dynamic range A-weighted (DR_A) [dB]	106
Schreier figure of merit (FoM_S) [dB]	180

components around f_s are aliased back to the audio band, but with an attenuation of more than 70 dB, in excess of the application requirements. This value is typical of a CT ΣΔM based on the CIFF topology.

The analog section of the third-order CT ΣΔM consumes 350 μW, while the digital blocks (i.e., DEM and thermometer-to-binary converter) consume 40 μW, both from a 1.6-V power supply and during conversion. The FoM_S is 180 dB. Table 2 shows a summary of the performance achieved by the CT ΣΔM.

5.2 *Future Trends*

Further efficiency improvements in microphone front-ends are under development, and some of them are reported here.

Higher Quantizer Resolution to Decrease Sensitivity to Clock Jitter
One major drawback of CT-ΣΔMs with respect to SC architectures is the increased DR degradation in the presence of clock jitter. In fact, in CT-ΣΔMs the jitter on the

clock used by the feedback DAC produces an equivalent noise component, which is directly added to the input signal, while this is not the case in SC structures, in which the clock jitter only affects the input signal sampling.

In first approximation [30], for a multibit CT-ΣΔM, the expected value of the signal-to-jitter-noise ratio (SJNR) is given by:

$$\mathrm{SJNR} = 10 \cdot \log_{10}\left[\frac{\left(2^N - 1\right)^2}{16 \cdot \mathrm{OSR} \cdot J_{\mathrm{RMS}}^2 \cdot B^2}\right] [\mathrm{dB}], \tag{3}$$

where J_{RMS} is the standard deviation of the clock jitter and N the number of bits of the quantizer. According to (3), a straightforward solution for reducing the performance degradation due to jitter is to increase the number of bits in the quantizer. However, if the quantizer is implemented with a conventional flash ADC, this would result in a more complex structure, larger power consumption, and larger silicon area.

Given the large OSR used for audio converters, tracking ADCs are a convenient solution to achieve high resolution while reducing power and area compared to classic flash ADCs, however, they can perform a proper conversion only if the input signal remains in the tracking range [31]. Wrong or missed conversions in a tracking ADC employed as quantizer in a ΣΔM ADC can lead to instabilities and oscillations.

In SC-ΣΔM, an anti-aliasing filter is required in the input path, and usually such filters are designed with a cut-off frequency just above the audio bandwidth. Therefore, if the tracking ADC can operate with a full-scale input signal at the cut-off frequency of the anti-aliasing filter, input signals at higher frequency will always stay in the tracking range since they are attenuated by the filter itself. In CT-ΣΔMs, the input signal is attenuated only by the loop-filter, which has a cut-off frequency one order of magnitude higher. A conventional tracking ADC, therefore, should be designed with a larger tracking range, thus increasing power consumption and area.

A solution to this problem can be a tracking ADC that is able to convert audio-band signals with full resolution, while performing only a coarse conversion when an input signal that exceeds the tracking range is applied, thus ensuring stability for the CT-ΣΔM.

The analysis of this solution can start referring to Fig. 26. It is worth noting that the sample-and-hold circuit (S&H) operates at the rising edge of clock Ck, while the feedback DAC is clocked at the rising edge of Ckn. Therefore, there is a delay time of half sampling period ($T_S/2$) in the feedback loop. Having such delay is a common solution in CT-ΣΔM, because it can relax the speed requirement of the quantizer.

A tracking ADC for a CT-ΣΔM is shown in Fig. 27. The number of comparators N_{tk} is a function of the final desired resolution of the tracking ADC (N_{ADC} levels), the audio bandwidth (B), and the sampling period (T_S). To a first approximation, N_{tk} is given by:

$$N_{\mathrm{tk}} = 2 \cdot \mathrm{round}\left[N_{\mathrm{ADC}} \cdot \pi \cdot B \cdot T_S\right] \tag{4}$$

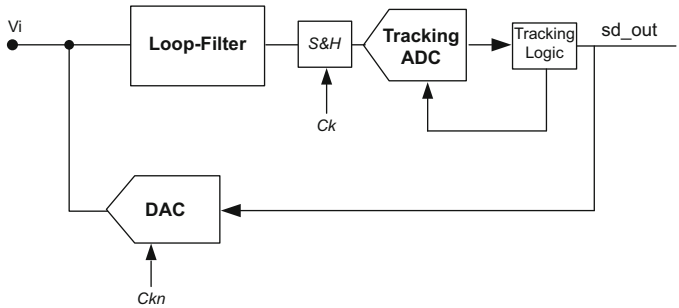
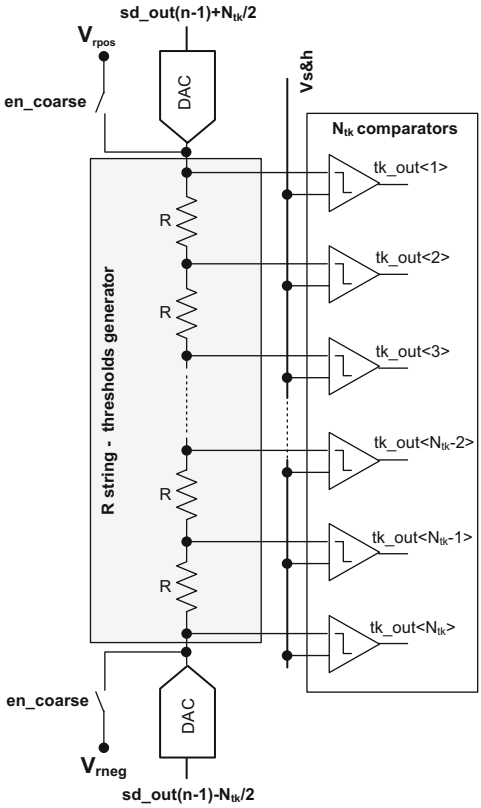


Fig. 26 CT-ΣΔM with tracking ADC

Fig. 27 Block diagram of the tracking ADC



The comparator thresholds can be generated with a resistor string. The voltage drop for each resistor R is equal to V_{FS}/N_{ADC}, where V_{FS} is the full-scale value of the signal to be converted.

The upper and lower ends of the resistor string are connected to two complementary DACs. Each DAC generates a voltage that is a function of the CT-ΣΔM's

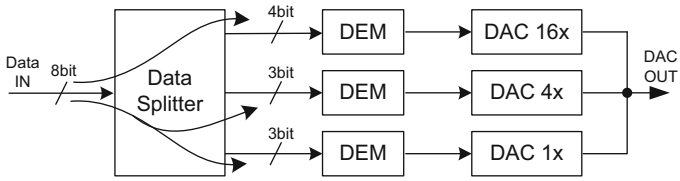


Fig. 28 Data segmentation for 8-bit DAC

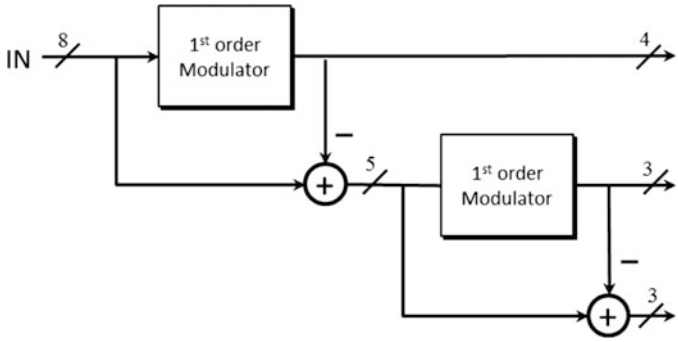


Fig. 29 Block diagram of an 8-bit 3-way data splitter

output previous, keeping the voltage drop across the resistive string constant and centered on the signal under conversion. The output of the CT-ΣΔM can thus be reconstructed from the previous conversion and the current output of the tracking ADC. If the tracking ADC output is at the limit of the tracking range (i.e., it is $+N_{tk}/2$ or $-N_{tk}/2$), a second coarse conversion is performed in the same conversion time window of $T_S/2$. The coarse conversion is performed by shorting the ends of the resistive string to V_{rneg} and V_{rpos}, where $V_{rpos} - V_{rneg} = V_{FS}$. If the result of this conversion is out of the tracking range, the Tracking Logic forces the use of coarse conversions in successive conversions, until the input signal returns in the tracking range.

Adaptive DEM in Feedback DAC

Increasing the number of quantizer bits has the drawback of increasing the complexity of the feedback DAC, particularly of the DEM logic. In order to reduce the DEM complexity, a technique known as segmentation (or noise-shaped splitting) can be used, in which the N bit digital signal at the output of the quantizer can be segmented in to multiple digital signals, each having less than N bits, so that each smaller segment can be processed and recombined with the other segments [32]. An example of this technique applied to an 8-bit digital signal is shown in Fig. 28. The data splitter can be realized as a cascade of first-order digital ΣΔMs, as shown in Fig. 29. There are two main drawbacks of this technique that limits the achievable DR. The first one is the effect of thermal noise, considering that the signal is processed by the DAC with the highest weight, while the DACs with

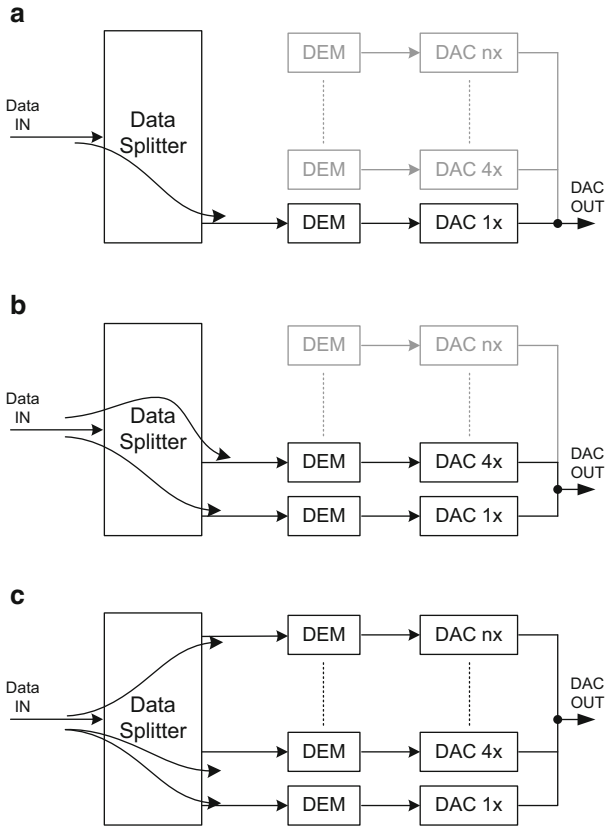


Fig. 30 Adaptive DEM

smaller weights are processing the quantization noise. Since the thermal noise is proportional to the weight of the DAC, the noise floor is dominated by the thermal noise generated by the DAC with the highest weight, even for small amplitude output signals, thus limiting the DR. The second drawback is the gain error between the DACs: the DAC-to-DAC error is shaped only by a first-order high-pass transfer function, thus again limiting the DR. Therefore, advanced layout techniques are required to minimize the mismatch between the DACs, increasing the complexity and the design area.

Another drawback is related to the power consumption and is correlated to the already mentioned fact that the signal is processed by the DAC with the highest weight: Even if the output signal is small, i.e., it is contained within few DAC levels, it is actually the result of the subtraction of a large signal generated by the DAC with the highest weight and the smaller quantization signals generated by the DACs with lower weights.

This means that all the DACs must be always active, i.e., the power consumption for small signal is comparable to the power consumption at full-scale. A solution to these problems is the use of an adaptive DEM scheme, in which [33], the segmented DAC can be dynamically reconfigured. An envelope detector tracks the amplitude of the digital signal at the input of the DAC. When the signal can be expressed with only the lowest-weight DAC (1×), the other segments are bypassed, and their DACs are turned off, as shown in Fig. 30a. Likewise, when the signal can be expressed only with the first- and second-lowest weight DACs (1× and 4×), the other segments are bypassed and their DACs are turned off, as shown in Fig. 30b. Finally, when the signal amplitude requires the DAC with the highest weight to be used, all the segments are turned on, as shown in Fig. 30c. The number of possible operational states is equal to the number of segments.

This solution overcomes several drawbacks of the previous technique. In small-signal operation (i.e., when only the DAC 1× is used), the thermal noise is lowered compared to large signal operation, increasing the DR. Moreover, the noise and distortion from DAC-to-DAC gain error is avoided, since only one DAC is used. Similar considerations can be made for the mid-level signal operation (i.e., when the segmentation is applied only to DACs 1× and 4×). Finally, a dynamic "Class-H"-like power consumption is achieved: for each operational state, the power consumption is given only by the DAC elements that are actually in use, while the other DAC elements can be turned off. This means that the power consumption is greatly reduced in the presence of small signals.

References

1. Hsu YC, et al. Issues in path toward integrated acoustic sensor system on chip. In: Proceedings of IEEE sensors; Lecce, Italy; 2008. p. 585–8.
2. Malcovati P, Maloberti F. Interface circuitry and microsystems. In: Korvink J, Paul O, editors. MEMS: a practical guide to design, analysis and applications. Dordrecht: Springer; 2005. p. 901–42.
3. Bajdechi O, Huijsing JH. A 1.8-V ΔΣ modulator interface for an electret microphone with on-chip reference. IEEE J Solid-State Circuits. 2002;37:279–85.
4. Chiang CT, Huang YC. A 14-bit oversampled delta-sigma modulator for silicon condenser microphones. In: Proceedings of IEEE IMTC; Singapore; 2009. p. 1055–8.
5. Pernici S, et al. Fully integrated voiceband codec in a standard digital CMOS technology. IEEE J Solid-State Circuits. 2004;39:1331–4.
6. van der Zwan EJ, Dijkmans EC. A 0.2-mW CMOS ΣΔ modulator for speech coding with 80 dB dynamic range. IEEE J Solid-State Circuits. 1996;31:1873–80.
7. Zare-Hoseini H, et al. A low-power continuous-time ΔΣ modulator for electret microphone applications. In: Proceedings of IEEE ASSCC; Beijing, China; 2010. p. 1–4.
8. Jawed SA, et al. A 828-mW 1.8-V 80-dB dynamic-range readout interface for a MEMS capacitive microphone. In: Proceedings of ESSCIRC; Edinburgh, UK; 2008. p. 442–5.
9. Picolli L, et al. A 1.0-mW, 71-dB SNDR, fourth-order ΣΔ interface circuit for MEMS microphones. Analog Integr Circuits Sig Process. 2011;66:223–33.
10. Le HB, et al. A regulator-free 84-dB DR audio-band ADC for compact digital microphones. In: Proceedings of IEEE ASSCC; Beijing, China; 2010. p. 1–4.

11. Citakovic J, et al. A compact CMOS MEMS microphone with 66-dB SNR. In: IEEE ISSCC digest of technical papers; San Francisco, USA; 2009. p. 350–1.
12. Weigold JW, et al. A MEMS condenser microphone for consumer applications. In: Proceedings of IEEE MEMS; Istanbul, Turkey; 2006. p. 86–9.
13. Scheeper PR, et al. A review of silicon microphones. Sensors Actuators A. 1994;44(1):1–11.
14. Bergqvist J, Gobet J. Capacitive microphone with a surface micromachined backplate using electroplating technology. J Microelectromech Syst. 1994;3(2):69–75.
15. Kasai T, et al. Novel concept for a MEMS microphone with dual channels for an ultrawide dynamic range. In: Proceedings of IEEE MEMS; Cancun, Mexico; 2011. p. 605–8.
16. Leinenbach C, et al. A new capacitive type MEMS microphone. In: Proceedings of IEEE MEMS; Wanchai, Hong Kong, China; 2010. p. 659–62.
17. Martin DT, et al. A micromachined dual-backplate capacitive microphone for aeroacoustic measurements. J Microelectromech Syst. 2007;16(6):1289–302.
18. Zou QB, et al. Design and fabrication of silicon condenser microphone using corrugated diaphragm technique. J Microelectromech Syst. 1996;5(3):197–204.
19. InvenSense Application Note AN-1003. Recommendations for mounting and connecting InvenSense MEMS microphones, Online.
20. Knowles Application Note AN-16. SiSonic design guide, Online.
21. Nicollini G, et al. A high-performance analog front-end 14-bit CODEC for 2.7-V digital cellular phones. IEEE J Solid-State Circuits. 1998;33:1158–67.
22. Barbieri A, Nicollini G. 100+ dB A-weighted SNR microphone preamplifier with on-chip decoupling capacitors. IEEE J Solid-State Circuits. 2012;47:2737–50.
23. Croce M, et al. Cap-less audio preamplifiers for silicon microphones. In: Proceedings of IEEE sensors, Orlando, FL, USA; 2016. p. 943–5.
24. Croce M, et al. MEMS microphone fully-integrated CMOS cap-less preamplifiers. In: Proceedings of IEEE PRIME, Giardini Naxos, Taormina, Italy; 2017. p. 37–40.
25. Jiang X, et al. A low-power, high-fidelity stereo audio CODEC in 0.13-μm CMOS. IEEE J Solid-State Circuits. 2012;47:1221–31.
26. Du D, Odame KM. A bandwidth-adaptive preamplifier. IEEE J Solid-State Circuits. 2013;48:2142–53.
27. Tsividis Y, et al. Internally varying analog circuits minimize power dissipation. IEEE Circuits Device Mag. 2003;19:63–72.
28. De Berti C, et al. A 106-dB A-weighted DR low-power continuous-time $\Sigma\Delta$ modulator for MEMS microphones. IEEE J. Solid-State Circuits. 2016;51:1607–18.
29. Murman B. ADC performance survey 1997–2017. Online. http://web.stanford.edu/~murmann/adcsurvey.html.
30. De Berti C, et al. Colored clock jitter model in audio continuous-time $\Sigma\Delta$ modulators. In: Proceedings of IEEE NEWCAS, Grenoble, France; 2015. p. 14B5/1–4.
31. Dörrer L, et al. A 3-mW 74-dB SNR 2-MHz continuous-time delta-sigma ADC with a tracking ADC quantizer in 0.13-μm CMOS. IEEE J Solid-State Circuits. 2005;40:2416–27.
32. Nguyen K, et al. A 108-dB SNR, 1.1-mW oversampling audio DAC with a three-level DEM technique. IEEE J Solid-State Circuits. 2008;43:2592–600.
33. Crespi L., et al. Audio digital-to-analog converter with enhanced dynamic range. US Patent Application No. 62/425,510, 2016.