# A 381.4 $\mu$W 114dB-A DR Companding VCO-ADC for MEMS Microphones using a Multirate Architecture

J. Granizo*, R. Garvi*, *Member, IEEE*, R. Carrero*, J. Fernandez *, V. Medina*, D. Straeussnigg†, A. Wiesbauer†, *Member, IEEE*, L. Hernandez*, *Senior Member, IEEE*

***Abstract*—High-performance digital MEMS microphones require ADC architectures that combine power efficiency with extended dynamic range. We propose here a MEMS digital readout ASIC that uses two VCOs in parallel with 12dB of gain difference and a frequency-to-digital conversion circuit that combines the VCO signals into a single digital output. As a main feature, the architecture benefits from VCO-ADC first order noise shaping and a large quantizer resolution of 11 bits to achieve a minimal glitch energy when the two VCO signals are switched. A demonstration 130nm CMOS chip shows a peak SNDR of 91dB-A, SFDR of 99dB and 114dB-A of dynamic range with open-loop VCOs in $0.136mm^2$. Total power is $381.4\mu W$ without requiring an extra buffer to connect a MEMS device.**

***Index Terms*—VCO-ADC, VCO-based ADC, Companding ADC, MEMS microphones, time encoding**

## I. Introduction

High dynamic range microphones are required for outdoor audio recording, noise cancellation and voice-based artificial intelligence. Due to that, the development of low noise and high acoustic overload MEMS microphones has become a priority for the industry. The industry standard for MEMS microphones is the switched-capacitor Sigma-Delta ADC [1], [2]. Lately, voltage-controlled oscillator based ADC, VCO-ADC, have emerged as an alternative to overcome the limitations of traditional Sigma-Delta ADC [3], [4], [5].

Two main architectures of VCO-ADC exist: the open-loop VCO-ADC and the closed-loop VCO-ADC, both of which exhibit first-order noise shaping [3]. Open-loop architectures do not need an input buffer to be MEMS compatible and present an acceptable linearity for microphones, making them more suitable for the audio application than closed-loop designs [6], [7]. Nevertheless, for high dynamic range and low noise microphones, the use of an open-loop VCO-ADC would lead to a prohibitively high power consumption.

An interesting option to extend the dynamic range of ADCs is the use of companding [8]. Fig 1 (a) illustrates the basic concept of companding. Depending on the input amplitude, the signal is amplified with a different gain prior to digitization with an ADC. For low input amplitudes, the gain is higher than for high input amplitudes, so that the SNR of the former is increased in the noise dominated region of the dynamic range (DR). Nevertheless, this scheme has a problem during transitions. In audio applications, high-order sigma-delta modulators are chosen as the best fit for the ADC in Fig 1 (a). Due to the memory of the Sigma-Delta modulator, a glitch appears when the gain of the signal-path is switched. To solve this problem, two different Sigma-Delta modulators can be used for each gain path, creating two different channels, as shown in Fig 1 (b). Although this solves the aforementioned problem of the glitch generation, the limited number of bits (1 - 5) of conventional Sigma-Delta modulators cause a discontinuity proportional to the LSB of their quantizer during transitions. This manifests as the in-band noise-folding shown in Fig 1 (b). As an advantage, audio VCO-ADCs have a quantizer with a higher number of bits (>9) [6] which results in a glitch energy proportional to a smaller LSB size. Therefore their use in a companding ADCs greatly mitigates the problem of noise-folding. This system is portrayed in Fig 1 (c).

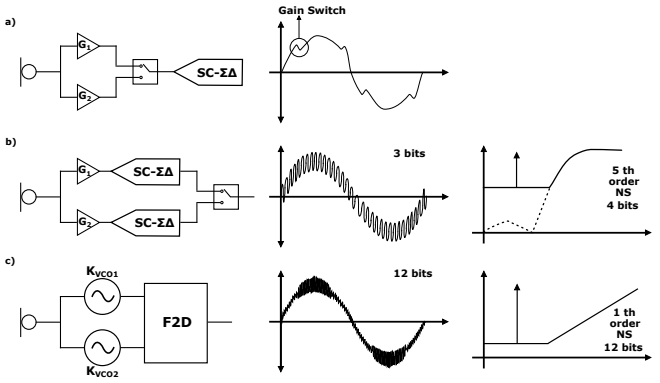


Fig. 1: (a) PGA-based companding architecture (b) Two-channel companding architecture. (c) VCO-ADC companding architecture

This paper presents a VCO-ADC for MEMS microphones fabricated in CMOS 130nm with companding structure. Furthermore, the use of a multirate noise shaper allows higher sampling frequency in the oscillator, lowering the oscillation frequency for better noise performance [9].

## II. System level description

Fig. 2 shows the system-level details of the proposed chip. The system is composed of two parallel signal paths (channels)

Research funded by projects PDC2023-145850-100 and FPU21/02257 of the Spanish Ministry of Science, Innovation and Universities

* Carlos III University, Leganes, Spain. V. Medina is now an Independent Researcher

† Infineon Technologies Austria Ag.

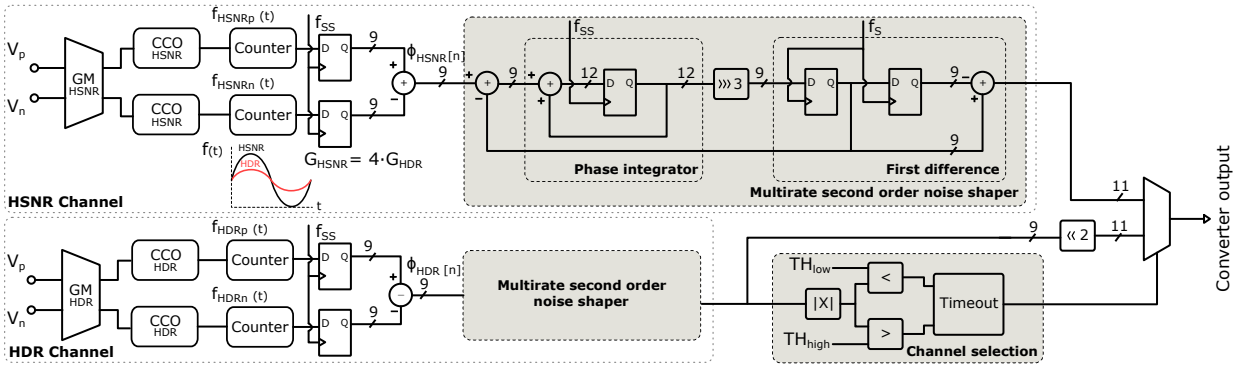


Fig. 2: System level view of the proposed VCO-ADC companding converter.

that process the same differential signal with different sensitivities. As a consequence, one of the channels (High Dynamic Range, HDR) is optimized to process the full dynamic range of the converter at the expense of a lower SNR. The second channel (High Signal-to-Noise-Ratio, HSNR) saturates at some point but guarantees a high sensitivity and low noise. The different sensitivities are implemented by both attenuating by 2 the HDR channel input and a gain factor of 2 in the HSNR channel, which yield a combined sensitivity difference of 4 (12dB). The output is selected among these channels by using an algorithm digitally implemented. Each of the channels is composed of a Current-Controlled Oscillator (CCO) followed by a counter implementing a Frequency-to-Digital converter (F2D). Then, the counter is sampled at $f_{ss} = 24MHz$ and processed with a decimate-by-8 multirate noise shaper clocked by $f_{ss}$ and $f_s = 3MHz$. The system outputs data at $f_s$. Inter-channel gain and offset correction is essential to mitigate the glitches produced in the channel transitions. The availability of both channel digital outputs simultaneously allows to easily find calibration parameters in the background. Moreover, two F2D blocks does not impose a significant penalty, as the power budget and area are dominated by the analog circuitry of the chip rather than the F2D blocks.

### A. Input stage

Each of the channels has a transconductor (GM) and differential ring CCO with 16 delay taps. In Fig. 2 it can be seen that the HSNR CCO frequency variation span is 4 times that of the HDR CCO. However, after the noise shaper, there is a digital gain compensation by 4 for the HDR channel. As a result, the signal path gain is the same regardless of the channel used. This factor of 4 is chosen because it provides a reasonable gain difference, and the digital correction can be done with a logic shift. The pulses in the 16 taps of the CCO, $\Phi(t)$ are integrated using a 9-bit counter [9] and sampled at 24 MHz. $\Phi_s[n]$ represents the sampled version of the counters, subtracting p and n branches of the differential analog channel.

### B. Noise shaper

A first-order noise-shaped signal can be obtained from sampled signal $\Phi_s[n]$ by a first-order difference. Such signal already encodes the input with the required SNR. However, industry standard microphones use sampling rates between 0.5 and 4MS/s but $\Phi_s[n]$ is sampled at 24 MHz. Thus, we have to use a rate converter to bring down the sampling rate to 3MHz. This rate converter is implemented as a multirate noise shaper similar to a Cascaded Integrator Comb (CIC) filter, shown in Fig. 2. This blocks keeps the same 9-bit word-length at the output by introducing second-order noise shaping of the decimation and truncation errors, while quantization error of $\Phi_s[n]$ remains first-order noise shaped. As a result, the noise shaper is composed of three blocks: a first integrator (CCO + counter), a digital first-order sigma delta modulator, and a first difference. Thus, the STF can be considered to be a $sinc^2$, and the NTF would be second-order noise-shaped. However, the SQNR limit is imposed by the input, in this case $\Phi_s[n]$. Since the HDR noise shaper receives a smaller signal (later digitally amplified), it will not saturate and thus both HSNR and HDR channels can use the same noise shaper block with a reduced number of bits. For comparison, using a single channel would require 2 more bits on the noise shaper as the analog input would see a single gain.

### C. Selection algorithm

Regarding the channel selection algorithm, it should address two primary issues. First, the channel switching must be done fast once the distortion of the input signal becomes significant, and second, it should prevent repeated channel switching in short time frames, which might result in audible glitches. Since in our implementation we have access to a sigma-delta modulated signal quantized with 9 bits, its time domain representation provides enough resolution to closely resemble the input analog signal. This allows for a fast measure the signal amplitude using digital comparators and without requiring a decimation filter (see "Channel selection" in Fig. 2). To prevent repeated, fast channel switching, we use a combination of two methods: hysteresis and timeout. Once the signal goes above $\text{TH}_{high}$, the output is immediately assigned to the HDR channel, preventing distortion. However, in this state, for the system to return to the HSNR channel, it is necessary that the signal amplitude stays below $\text{TH}_{low}$ for a programmable period of time. This results in the system always switching channels at a frequency in the subsonic region, and thus no audible tone should result from its activity.

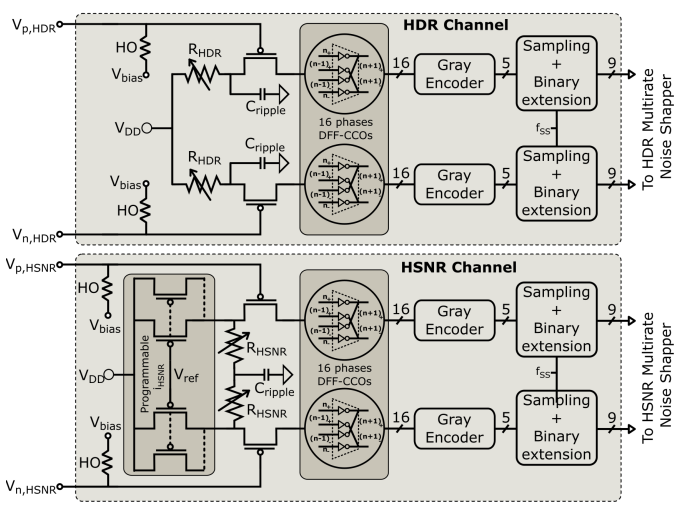


Fig. 3: HSNR and HDR Channel analog input stages

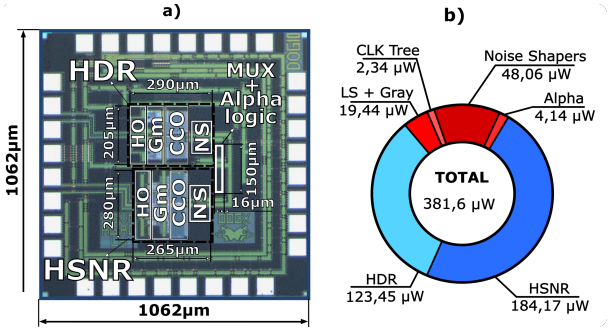


Fig. 4: (a) Die micrograph. (b) Power breakdown, analog (1.5V) corresponds to blue sectors, while digital (0.9V) corresponds to red sectors.

## III. Circuit design

Both HSNR and HDR channels make use of a degenerated-source transconductor (GM) to modulate the CCO, as shown in Fig. 3. In the case of the HSNR GM, a differential structure with a current mirror powering up each CCO is used. This ensures that the total current used by the HSNR CCOs remains constant, even at amplitudes exceeding the HSNR GM expected DR. In the HDR channel, a pseudodifferential structure with no current mirror is employed, which helps improve the linearity of the GM at high amplitudes. PVT compensation of the oscillation frequency is obtained by means of a programmable current mirror in the HSNR channel, and an external biasing voltage in the HDR channel. In both transconductors, $C_{ripple}$ is placed to remove oscillator switching coupling between the p and n branches. For PVT gain variations and to ensure a difference of 4 times the gain between HSNR and HDR channels, 4-bit programmable resistors are used. The resolution of the HDR programmable resistor allows for 0.4dB gain steps, resulting in a maximum 0.2dB gain difference between the channels. This difference can be further compensated in the digital logic to ensure seamless switching. The CCO is a 16 tap differential feed-forward (DFF) ring oscillator, based on the design proposed in [10]. This oscillator allows an even number of taps while presenting similar noise efficiency to a single-ended oscillator and better linearity. Using both the positive and negative terminals of each differential tap, oscillation phases are converted to the necessary digital logic levels using level shifters.

Regarding the 9-bit counter commented on Fig 2, the first stage consists of a phase-to-Gray encoder [9], which transforms the 32 possible states of the oscillator into a 5-bit Gray sequence. This sequence is then sampled at $f_{ss}$, and extended utilizing an asynchronous counter into 9 bits. This output is then used by the multirate noise shaper, and downsampled to the required 3MHz. The encoding of the state of the oscillator into a Gray sequence before sampling helps reduce the required number of D flip-flops (from N taps to $\log_2(N)$), converting them into more area and power efficient XOR gates, and avoiding the thermometric-to-binary encoder. Furthermore, the use of Gray encoding means that a metastability error given by the sampler would result in only 1 LSB of quantization error.

## IV. Implementation and Measurements

The proposed companding VCO-ADC has been implemented in a 130 nm CMOS process, occupying an active area of 0.136 $mm^2$, as seen in Fig 4. The total power consumption of the converter is 381.6 $\mu$W at 1.5V in the analog supply and 0.9V in the digital supply. The chip has been tested with an audio generator and a quartz clock together with a logic analyzer and a micro controller board for configuration. In our measurement setup, an additional 6dB attenuator is placed between the input signal generator and the HDR channel. Although placed externally, it can also be placed internally within the ASIC, employing a capacitive divider [8]. PSDs featuring the peak SNDR of the HSNR and HDR channel are shown in Fig. 5. The use of a differential feed-forward oscillator in conjunction with the pseudodifferential GM input stage results in a very linear current-to-frequency curve, despite not having any feedback loop. The noise-shaping $1^{st}$ order and $2^{nd}$ order slopes, given by the noise shaper, are also appreciated in both spectra.

Fig 6 shows the dynamic range (DR) of the full ADC. The full scale of the converter sits at 4dBV differential input. This corresponds to a -36dBV signal at -40dBFS, which is considered a typical sensitivity of a microphone MEMS at 1Pa/94dB$_{SPL}$ [6], [8], [9]. SNDR of the converter at -36dBV is 74.3dB-A. As in previous works, total DR in a microphone ADC ASICs is considered at the AOP point. In our case, the AOP is considered at 5% THD, which occurs at 0dBFS. As such, the DR of the ADC with both channels active is 114dB-A. Note that at low input signal power, the HSNR channel is the one at the output of the converter. Once the input signal power is sufficient to generate distortion in the HSNR channel, at -32dBFS, HDR channel is selected as the output. Fig 6 also plots the DR of the oscillators at the HSNR and HDR channels.

The output of the ADC against a dual tone input, one at 1Hz -20dBFS and another at 1kHz -40dBFS, is shown in Fig. 7. The displayed output has been digitally calibrated off-chip to ensure no gain or offset mismatch and normalized between

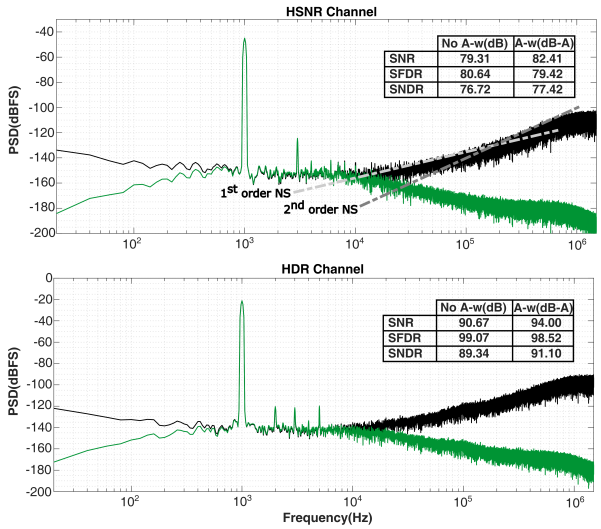


Fig. 5: PSD of peak SNDRs in HSNR (-26dBFS) and HDR (-10dBFS) channels. Green Line shows FFT with A-weighting.

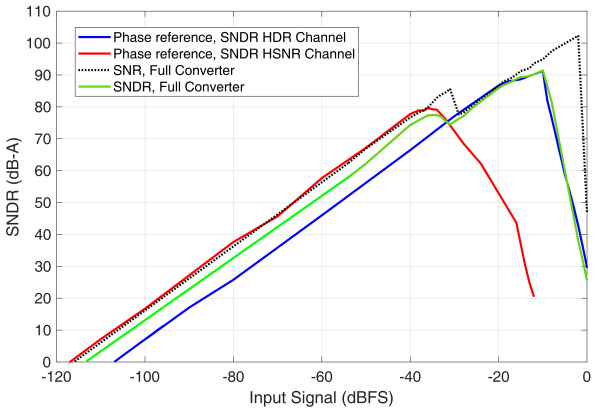


Fig. 6: DR of full ADC, $TH_{high}$ threshold at -32dBFS.

-1 and 1. The high-pass filter to remove the subsonic tone and harmonics is placed at 10Hz. The boundaries between HSNR output and HDR output of the ADC can be easily seen as the quantization step size changes. At low amplitudes, HSNR channel is the output and the quantization steps are 1LSB, while at high amplitudes, HDR channel takes over, resulting in 4LSB quantization steps. Nevertheless, transition glitches between HSNR and HDR output are kept below -60dB with respect to the audible tone (1kHz), making them inaudible.

## V. Conclusions

Table I shows our work against other state-of-the-art ADC implementations with similar bandwidths. MEMS microphones ASICs are reported using A-weighting filter. Our work shows competitive FoMs considering the effective DR for both the whole converter and also for the HDR channel considered as a single ADC. As a conclusion, we have proven that VCO-ADCs are an advantageous choice for companding architectures.

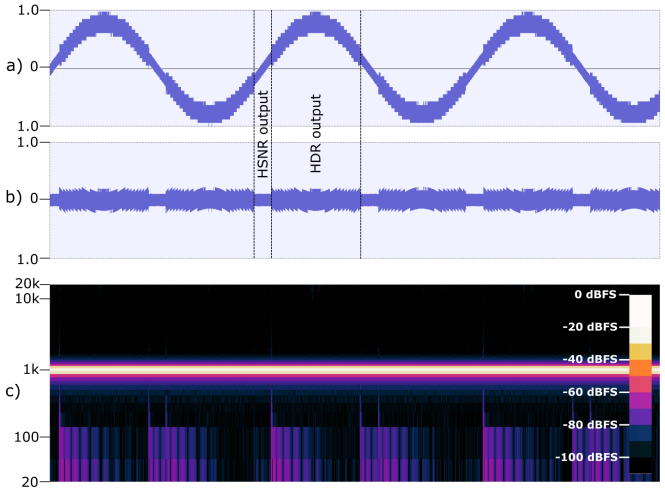


Fig. 7: (a) Dual tone input. (b) High-pass filtered output. (c) Spectrogram.

TABLE I: Comparison with other works

| | **This work** | [6] Sensors 2023 | [8] ISCAS 2024 | [1] JSSCC 2020 | [2] CICC 2024 | [5] ASSCC 2022 |
|---|---|---|---|---|---|---|
| Process (nm) | **130** | 130 | 130 | 180 | 180 | 28 |
| Topology | **Companding VCO-ADC** | Open-loop VCO-ADC | DTDSM | MASH CTDSM | Nested CTDSM | Closed-loop VCO-ADC |
| Input | **Hi-Z** | Hi-Z | Hi-Z | Low-Z | Low-Z | Hi-Z |
| BW [kHz] | **20** | 20 | 20 | 24 | 20 | 10 |
| Area [mm$^2$] | **0.136** | 0.14 | 1.1 | 0.64 | 0.74 | 0.095 |
| Supply (A/D) [V] | **1.5/0.9** | 1.5/0.95 | 1.8/0.9 | 1.8 | 1.8 | 1.1/0.55 |
| Power [$\mu W$] | **381.6** | 438.1 | 460 | 256 | 470 | 7.1 |
| $F_s$ [MS/s] | **3** | 3.072 | 3 | 6.144 | 7.68 | 0.72 |
| $SNDR_{peak}$ [dB],[dB-A] | **91.1*** | 80.31* | 92** | 100.9 | 107.3 | 78.6 |
| $SFDR_{peak}$ [dBc] | **99.1** | | | 108 | 121.6 | 95.3 |
| DR [dB],[dB-A] | **114.3*** | 108* | 108* | 104 | 109.2 | 79.3 |
| $FoM_{SNDR}$ [dB],[dB-A] | **168.3*** | 157* | 168.3** | 180.5 | 183.6 | 170.1 |
| $FoM_{DR}$ [dB],[dB-A] | **191.2* (HSNR) 179.9* (HDR) 186.0*** | 184.6* | 184.4* | 183.6 | 185.5 | 170.8 |

* Use of A-weighting filter. ** Estimated from DR, with A-weighting filter.

## References


[1] S. Billa *et al.*, "Analysis and Design of an Audio Continuous-Time 1-X FIR-MASH Delta-Sigma Modulator," *IEEE JSSC*, vol. 55, no. 10, pp. 2649–2659, 2020.

[2] J. Jin *et al.*, "A 470$\mu$W 20kHz-BW 107.3dB-SNDR Nested CT DSM Employing Negative-R-Based Cross-RC Filter and Weighted Multi-Threshold MSB-Pass Quantizer," in *IEEE CICC*, 2024, pp. 1–2.

[3] G. Gielen *et al.*, "Time-Encoding Analog-to-Digital Converters: Bridging the Analog Gap to Advanced Digital CMOS-Part 1: Basic Principles," vol. 12, no. 2, pp. 47–55.

[4] E. Gutierrez *et al.*, "Why and How VCO-based ADCs can improve instrumentation applications," in *IEEE ICECS*, 2018, pp. 101–104.

[5] Y. Zhong *et al.*, "A 78.6 dB-SNDR 520mVpp-full-scale 620Mohm - Zin 105dBCMRR VCO-based Sensor Readout Circuit Using FVF-Based Gm-Input Structure," in *IEEE A-SSCC*, pp. 1–3.

[6] C. Perez *et al.*, "A VCO-Based ADC With Direct Connection to a Microphone MEMS, 80-dB Peak SNDR and 438-$\mu$W Power Consumption," *IEEE Sensors Journal*, vol. 23, no. 8, pp. 8466–8477, 2023.

[7] C. Perez *et al.*, "A 73dB-A Audio VCO-ADC Based on a Maximum Length Sequence Generator in 130nm CMOS," *IEEE TCAS-II*, vol. 68, no. 10, pp. 3194–3198, 2021.

[8] J. L. Ceballos *et al.*, "A 70dBA-460$\mu$W Companding Digital Silicon Microphone with Programmable Acoustic Overload Point and MEMS Asymmetry Robustness," in *IEEE ISCAS*, 2024, pp. 1–5.

[9] V. Medina *et al.*, "Second-Order VCO-ADC Architecture With Low-Area and High Dynamic Range Using Internal Binary Encoding," *IEEE TCAS-I*, pp. 1–12, 2025.

[10] J. Borgmans *et al.*, "The Analog Behavior of Pseudo Digital Ring Oscillators Used in VCO ADCs," *IEEE TCAS-I*, vol. 68, no. 7, pp. 2827–2840, 2021.