Continuous-Time Acquisition of Biosignals Using a Charge-Based ADC Topology

Michal Maslik, Student Member, IEEE, Yan Liu, Member, IEEE, Tor Sverre Lande, Fellow, IEEE, and Timothy G. Constandinou, Senior Member, IEEE

Abstract—This paper investigates continuous-time (CT) signal acquisition as an activity-dependent and nonuniform sampling alternative to conventional fixed-rate digitisation. We demonstrate the applicability to biosignal representation by quantifying the achievable bandwidth saving by nonuniform quantisation to commonly recorded biological signal fragments allowing a compression ratio of ≈5 and 26 when applied to electrocardiogram and extracellular action potential signals, respectively. We describe several desirable properties of CT sampling, including bandwidth reduction, elimination/reduction of quantisation error, and describe its impact on aliasing. This is followed by demonstration of a resource-efficient hardware implementation. We propose a novel circuit topology for a charge-based CT analogue-to-digital converter that has been optimized for the acquisition of neural signals. This has been implemented in a commercially available 0.35 μm CMOS technology occupying a compact footprint of 0.12 mm². Silicon verified measurements demonstrate an 8-bit resolution and a 4 kHz bandwidth with static power consumption of 3.75 μW from a 1.5 V supply. The dynamic power dissipation is completely activity-dependent, requiring 1.39 pJ energy per conversion.

Index Terms—Analog-to-Digital Converter, Biosignals, Continuous-Time, CT-ADC, EAP, ECG, LFP.

I. INTRODUCTION

In recent decades we have seen tremendous interest in the development of electronic devices that record, process and communicate biological signals such as Electrocardiograms (ECG), Electroencephalograms (EEG), Local Field Potentials (LFP) and Extracellular Action Potentials (EAP). These are used in a variety of applications including medical devices (for diagnosis, therapy, and early detection) and tools for scientific investigation (for example, electrophysiology for neuroscience). Most such systems [1]–[3] operate by amplifying the weak biosignals, conditioning and filtering, followed by digitisation using uniform (i.e. fixed rate) sampling.

Many of the recorded signals (especially ECG [4] and EAPs) however exhibit short active periods in the form of spikes interrupting potentially long silent periods. This can lead to significant oversampling if these signals are sampled uniformly.

As there is currently a significant drive to create autonomous wirelessly powered implants [2], [5] it is necessary to ensure as power-efficient operation of all circuits as possible and it is therefore desirable to employ an alternative sampling scheme that would lead to sampling rate following the instantaneous spectral content of the signal of interest.

An example of such can be an adaptive sampling scheme increasing the sampling rate whenever activity is detected in the signal [6]–[8]. Such an approach can be effective but is dependent on the ability of detecting higher spectral content in the signal which typically proves to be difficult especially in low signal-to-noise-ratio (SNR) scenarios.

Another approach leading to bandwidth savings in transmission of sparse signals is Compressed Sensing (CS) allowing representation and reconstruction of certain signals from a smaller amount of samples than dictated by the Nyquist rate. Even though several implementations of complete CS-based systems recording biological signals [9], [10] have been demonstrated in recent past, those are often limited by the need of off-chip calculation of the used sparsifying matrix and increased power consumption due to a requirement of a digital-to-analog converter (DAC) operating at Nyquist rate.

A novel alternative is the use of Continuous-Time (CT) sampling [11], also known as Level-Crossing or Event-Driven sampling. As seen in Fig. 1 rather than taking samples at specific time periods, a CT Analogue-to-Digital Converter (CT-ADC) only takes samples when the signal crosses quantisation levels. This inherently leads to the sampling rate being dependent on the instantaneous spectral content making this approach ideally suited for digitisation of biological signals which are sparse in nature. This relatively new form of data conversion also presents new opportunities for new signal processing paradigms such as clockless CT digital filters [11], [12] or spike sorting [13].

Over recent years there have been several demonstrated implementations of CT-ADCs shown in literature [14]–[22]. Those are typically based on two comparators comparing the input signal to the two closest quantisation levels V_{up}^{ref} and V_{down}^{ref} that are either generated by a complete N-bit DAC or otherwise as summarised in Section III. In the event of the input crossing...
either of these levels, an output in the form of a pulse is generated leading to biphasic pulse output coding as described in [23].

In this paper which is an extension of work that has been presented at ISCAS 2017 conference [27] we present an alternative approach to CT-ADC front-end implementation making use of only a single comparator alleviating the need for an N-bit DAC by employing a charge-based solution reconstructing the quantised signal as a potential across a capacitor. The generation of capacitor potential is ensured by the use of two pulse generators adding or removing a precise amount of charge to or from the capacitor resulting in one Least Significant Bit (LSB) step.

The remainder of this paper is organised as follows: Section II presents an analysis and specific features of CT sampling with the aim of showing its suitability for acquisition of certain biological signals; Section III describes circuit-level implementation of the proposed CT-ADC topology; Section IV presents the silicon-verified measurements; and Section V concludes the paper.

II. CT Sampling for Biosignals

CT sampling has several desirable properties when compared to standard uniform sampling. One of its greatest advantages is the fact that samples are ideally taken exactly at points when the signal crosses quantisation levels and hence the exact value of each sample is known. This leads to elimination of quantisation error in each sample and can therefore lead to great improvements of reconstructed signal’s signal-to-noise and distortion ratio (SNDR) even allowing perfect reconstruction if the sampler is otherwise ideal [11].

A. Nonuniform Sampling

As is the case of any scheme involving non-uniform sampling, CT sampled signal can be reconstructed without aliasing if the average sampling rate is greater than the Nyquist rate [28] defined as twice the bandwidth or highest frequency content of the acquired bandlimited signal. This extension of the traditional Nyquist theorem however does not consider signals that change in spectral content over time such as biological signals but merely considers the overall bandwidth without any temporal resolution.

B. Variable-Bandwidth Signals

Since it is our aim to reduce the amount of taken samples in order to pave a way to reduction of transmitted data and hence power consumption it is necessary to consider sub-Nyquist sampling. As shown by K. Horiuchi [29] it is theoretically possible to devise a sampling scheme following the changes of spectrum of the acquired signal leading to a sampling scheme more efficient than uniform sampling requiring sampling rate higher than the Nyquist rate.

It is typically only possible to devise sampling schemes more efficient than Nyquist sampling if the signal of interest is sparse in some domain. A possible way of exploring sparsity of signals is treating bandwidth as a time-varying function and introducing the notion of average bandwidth. We thus aim to obtain a function $\omega_{\text{avg}}(t)$ representing the instantaneous spectral content of the given signal. As shown by [30] one possible approach to this problem is obtaining a function transforming time $t \rightarrow w(t)$ such that $s(w(t))$ becomes a constant-bandwidth version of signal $s(t)$. It then intuitively follows that more time contraction should be present at periods of higher spectral content and therefore $dw(t)/dt$ must represent an estimator of $\omega_{\text{avg}}(t)$.

C. Average Bandwidth Estimation

An alternative approach is obtaining a representation of the signal’s spectrum by applying a transform having a temporal resolution property such as a Short-Time Fourier Transform (STFT) [30]. STFT however suffers from fixed frequency resolution which limits its performance when used for bandwidth estimation. To overcome this limitation, the analysis presented in this paper is based on continuous wavelet transform $S_w(a, b)$ of signal $s(t)$ using Morse wavelets [31] of symmetry and decay parameters $\gamma = 1, \beta = 60$.

The instantaneous bandwidth of $s(t)$ is then estimated by thresholding the wavelet transform such that the bandwidth of the signal is assumed to be the corresponding frequency of the highest-scale bin of $S_w(a, b)$ that contains energy $|S_w(a, b)|^2$ larger than threshold $T_N$. This is defined as $T_N = \mu_s^2 + 3 \times \sigma_s^2$ where $\mu_s$ and $\sigma_s^2$ represent the mean of the energy spectrum $|S_w(a, b)|^2$ and variance of $S_w(a, b)$ respectively making $T_N$ act as a noise floor estimator. This is in line with techniques that have been demonstrated for use in denoising of biological signals [32] by removing spectral content not occupying signal bandwidth estimated using wavelet transforms.

As our aim is to explore the possibility of compressing various biological signal such as ECG, EEG, LFP and EAP, the wavelet bandwidth estimation method has been applied to chosen fragments of those signals. The result obtained in Fig. 2 shows used time-domain signals alongside estimates of their bandwidth variations over time. A compression ratio $C = f_{\text{max}}/f_{\text{avg}}$ of maximal bandwidth $f_{\text{max}}$ and average bandwidth $f_{\text{avg}}$, calculated from the obtained results and summarised in Table I can be used as an indicator of potential bandwidth savings if appropriate non-uniform sampling is used to capture each signal of interest.
Fig. 2. Demonstration of changes in bandwidth of various biological signals. Left to Right: ECG, EEG, LFP, EAP. Top: Portion of time-domain signals used for analysis. Bottom: Spectral content over time obtained by wavelet analysis of the time-domain signal and thresholding. Dashed line represents average bandwidth. LFP signal is obtained by applying a high-order low-pass filter of a 300 Hz cut-off to a simulated neural signal. EAP signal is obtained by applying a high-order high-pass filter of a 300 Hz cut-off to a simulated neural signal. ECG and EEG fragments are sourced from [24] and [25] respectively and obtained using [26].

| Table I: Summary of Biological Signals’ Bandwidth |
|-----------------------------------------------|----------------|----------|---------|---------|
| ECG | EEG | LFP | EAP |
| Maximal Bandwidth $f_{max}$ (Hz) | 51.07 | 40.24 | 357.52 | 7548.09 |
| Average Bandwidth $f_{avg}$ (Hz) | 10.97 | 7.32 | 52.38 | 289.46 |
| Compression Ratio C | 4.65 | 5.49 | 6.8 | 26.3 |

D. ECG Characteristics

It can be seen that ECG signals contain periodically repeating intervals of high spectral content concentrated around each QRS complex. Using an ideal sampling scheme a compression ratio close to 5 could in theory be achieved without loss of information. This can be compared to ECG compression algorithms reported in literature [33] typically achieving compression ratios around 2.5 in lossless operation. Due to their well-described sparseness, it is generally understood that ECG signals are a prime example of signals suitable for CT acquisition which is demonstrated by complete wireless ECG recording nodes demonstrated in literature such as [34].

E. EEG and LFP Characteristics

Even though the theoretical compression ratio achievable during quantisation of EEG or LFP signals is below 7 it should be noted that those signals typically have $1/f$ frequency characteristic [35] and therefore the used thresholding method might not be appropriate as it might neglect higher-frequency, lower-amplitude signals that typically carry a large amount of information. Similarly, there is not much reason to believe that those signals would benefit from variable sampling rates as, unlike the case of ECG and EAP, most components of those signals are present at all time and do not have bursty behaviour. In addition, it should be noted that apparent spikes of activity in the recorded LFP signals are caused by low-frequency remnants of EAP signals and do not necessarily indicate a property of LFPs.

F. EAP Characteristics

On the other hand, EAP signals are typically quiet and interrupted by sporadic spikes of high spectral content. As seen in Table I using adaptive sampling can lead to significant savings as the achievable compression reaches almost 27. This is in line with recent publications demonstrating compression ratios around 30 [36] using compressed sensing techniques.

G. CT Sampling Adaptive Properties

Having established that EAP signals possess properties making them suitable for the use of variable sampling rate acquisition it is possible to demonstrate that CT sampling leads to local-bandwidth-dependent sampling rate and is therefore a suitable technique for acquisition of neural spikes. To demonstrate this a simulation has been carried out resampling finely quantised EAP signal using CT sampling. The quantisation level size was set such that the signal crosses 256 levels and the largest amplitude portions of the signal are therefore quantised with the equivalent of 8 bits. As seen in Fig. 3 the sampling rate closely follows the local bandwidth obtained using wavelet analysis and demonstrates that CT sampling shows properties of bandwidth-dependent sampling.
The average achieved sampling rate is $f_s = 25.9$ kHz. This being significantly larger than expected is due to the fact that using CT sampling in a scenario with more than 1 quantisation level necessarily leads to oversampling. This can be demonstrated on an example of a pure tone as demonstrated in Fig. 4. Assuming the input signal is a single sinewave of arbitrary frequency and is CT-sampled with a resolution of N bits, it has to pass $2 \times 2^N$ quantisation levels during each cycle. Bearing in mind that only two samples per cycle are needed for reconstruction, the signal is oversampled $2^N$ times. Using the estimate obtained in Table I we find that the actual achieved oversampling ratio is close to 90. The fact that this is comparably smaller than the expected value of 256 is to large extent due to the fact that the input signal does not occupy the full input dynamic range most of the time as seen in Fig. 2.

Even though increasing the sampling resolution in CT sampling leads to an increase of oversampling, it is worth noting that if the system was ideal, it does not in theory lead to a decrease of quantisation error as an ideally CT-sampled signal should always be quantisation-error-free. This would suggest that it is optimal to always use single-level quantisation rather than multiple levels. However, in real-life implementations, quantisation is subject to errors introduced in the acquisition circuit and increasing the amount of quantisation levels can lead to improved performance.

H. CT Sampling and Aliasing

It could be expected that CT sampled signal is not prone to aliasing as the sampling rate directly follows the spectral content of the signal ensuring that it exceeds the Nyquist rate at each time instant if certain conditions are met. It is however important to understand that aliasing can still occur due to limited sensitivity inherent to the used amplitude quantisation. This imposes limitations on the input in terms of amplitude rather than bandwidth as is typical for uniform sampling.

If we consider CT sampling process where the size of a quantisation level is $\Delta LSB$ it is obvious that an input signal might not be recorded at all if its peak-to-peak amplitude is smaller than $\Delta LSB$ which gives rise to the sampler’s sensitivity. However, even if the peak-to-peak amplitude is larger than $\Delta LSB$, it is still possible that aliasing is going to occur if the input signal does not pass enough quantisation levels as to make the sampling rate exceed the Nyquist rate. Finding a sufficient condition on the input signal that would guarantee reconstructability is generally considered extremely difficult [37] as it would require solving the problem of obtaining probability density of generic signal quantisation level crossings. While this has been to some extent solved for certain classes of signals, such as Gaussian processes [38] a generic solution still remains undiscovered. It is however possible to express at least a minimal condition that guarantees perfect reconstruction of the input if satisfied:

**Proposition 1:** CT sampled signal can be ideally reconstructed if at any time instant the original signal is bandwidth limited to a bandwidth of $B$ and the component at frequency $f$ has a peak-to-peak amplitude of at least $\Delta LSB$ where $\Delta LSB$ is the step between two closest quantisation levels.

If this condition is satisfied it is ensured that the input generates at least $2B$ samples per second and hence can be reconstructed without any loss. This then applies to any component of the signal even if its peak-to-peak magnitude is smaller than the sensitivity of the system.

This is illustrated by Fig. 5 in time domain and Fig. 6 in frequency domain showing effects of aliasing in CT sampling due to quantisation of signals smaller than sensitivity of the sampler. Aliasing images are marked by dashed lines.

$$y_1(t) = \sin(0.7 \times 2\pi t) + 0.4 \sin(0.2 \times 2\pi t)$$

$$y_2(t) = 0.4 \sin(0.7 \times 2\pi t) + \sin(0.2 \times 2\pi t)$$
quantised by a CT sampling process using two quantisation levels of $\Delta LSB \approx 1.4$ giving rise to sampled signals $y_1(t)$ and $y_2(t)$. While in $y_1(t)$ the higher frequency component has peak-to-peak amplitude larger than the sensitivity of the acquisition process and the smaller frequency component has amplitude smaller than the sensitivity, the opposite applies for $y_2(t)$. The sampled signals are reconstructed using sinc interpolation method described in [39] giving rise to $y_1(t)$ and $y_2(t)$ respectively. The used method guarantees perfect reconstruction if the extended Nyquist condition is satisfied.

It can clearly be seen that reconstruction of $y_1(t)$ is possible while that of $y_2(t)$ is not. This is as $y_1(t)$ satisfies the condition given by Proposition 1 whereas $y_2(t)$ does not. As seen in Fig. 6 the insufficient average sampling rate occurring during quantisation of $y_2(t)$ leads to creation of an alias of the sub-sensitivity high-frequency component in the band of the signal preventing reconstruction. The SNDR of $y_1(t)$ after reconstruction defined as

$$SNDR = 10\log\left(\frac{E(y_1(t))}{E(|y_1(t) - y_1(t)|)}\right)$$

where $E(\cdot)$ marks the energy of each signal was found to be 115.5 dB which vastly exceeds SNDR of $\approx 7.78$ dB that would otherwise be achievable using discrete-time 1-bit quantisation.

Similarly to the issue of obtaining a minimal condition guaranteeing reconstructability, it is equally difficult to obtain a precise expression of the aliased version of the signal. This is due to the fact that knowledge of the average sampling rate would be required. That is however difficult to obtain as argued earlier.

It can be empirically seen that changing the size of a quantisation level has an effect on the sampling rate and hence the image frequency. Typically, reducing the quantisation level step leads to a higher sampling rate as expected.

I. Effect of Time Quantisation

Even though it is in principle possible to build complete systems such as filters [11], [40] or neural spike processors [13] that entirely operate in continuous time and provide real-time output without any quantisation, it is necessary to quantise the obtained time information if the acquired signal is to be digitally stored. This is inherently leading to introduction of in-band quantisation error [11] that results in reduced SNR that as shown in [20] can be expressed in a similar fashion to the 6 dB law giving a relationship between amplitude quantisation and maximal achievable SNR in discrete-time systems. In terms of the maximal frequency content of the signal $f_{max}$ and frequency of timer clock $f_{min}$ the maximal achievable SNR of the system is then:

$$SNR^{MAX} = 20 \times \log\left(\frac{f_{min}}{f_{max}}\right) - 11.2 \text{ dB}$$

This demonstrates, similarly to observations presented in [11] that infinitely fine quantisation equivalent to truly CT operation leads to no wide-band noise in the quantised signal. The effect of quantisation itself is generation of harmonics that can in theory be completely removed if an ideal reconstruction scheme is used, thus leading to a theoretically infinite SNDR. [11]

In addition, (4) shows that transmission of CT-sampled signal would benefit from adaptive clock frequency that would increase with an increase in signal spectral content indicated by an increased sampling rate. This could lead to a devision of a more efficient digitisation circuit or transmission scheme.

I. Analogue Storage of CT Signals

Even if signal storage is required, it is still possible to completely remove the need for time quantisation and create a fully CT system. While short-term storage of continuous time information can be achieved using a transistor-based delay cell [11], long-term non-volatile storage would in theory be possible e.g. on magnetic or optic media. Furthermore, it is possible to implement a CMOS-based non-volatile storage using memory cells such as those described in [41] paving a way to completely clockless CMOS-based recording circuits.

III. PROPOSED CT-ADC TOPOLOGY

Even though there have already been several presented implementations of CT ADCs designed for use in acquisition of ECG [17]–[19] or neural signals [14]–[16], [20], [21], a great challenge of those topologies remains the design of a feedback loop generating the reference levels following the input signal. As the traditional approach requiring the use of a complete N-bit feedback DAC [14], [17], [18], [20] leads to an increase in complexity and power consumption, several alternatives have been demonstrated in the literature.

These include the use of a capacitor-based offset injection circuit [15], [16], or using a sample-and-hold circuit [19], [21] to sample the input signal, calculate an error signal (within the range of two quantisation levels) and compare it to fixed reference voltages. A recent implementation [22] features a fully-differential input integrator switching its polarity each time a quantisation level is crossed and thus folding the input signal to a fixed window similarly allowing use of a fixed reference voltage.

These designs will show additional power consumption due to a need for a complete DAC or two comparators potentially consuming large amounts of static power. As an alternative alleviating these issues we are therefore proposing an architecture illustrated in Fig. 7. This only uses a single amplifier and does not require a feedback DAC.

A. High-level Operation Description

The proposed circuit operates as follows (demonstrated by simulated sample waveforms seen in Fig. 8): The last quantised sample $V^-_{n}$ is held as a potential across a balancing capacitor $C_b$. This is connected to an inverting input of a differential amplifier comparing it to input signal $V_{in}$. This arrangement allows use of a single amplifier acting both as a comparator as well as an input amplifier if designed for high-enough gain.

The output of the comparator, $V_{comp}$ represents difference between the last quantised sample $V^-_{n}$ and the current input $V_{in}$. This is fed to a threshold-crossing detector composed of inverters with varying switching points indicating when one of
the quantisation levels was crossed. Once that happens a pulse generator is triggered generating a pulse indicating a shift up or down in quantisation levels. This acts as an output while at the same time controlling a charge pump adding or removing certain amount of charge from capacitor \(C_b\). If the amount of charge \(\Delta Q\) is such that it results in voltage difference \(\Delta V_b = \Delta Q / C_b\) where \(\Delta V_{\text{LSB}}\) is equal exactly to one Least Significant Bit (LSB), the potential at \(C_b\) is now again equal to the quantised version of the input and \(V_{\text{comp}}\) returns to its steady state. The cycle can then repeat for another threshold crossing and another sample generation.

As seen in Fig. 7 the implemented design is complemented by a configuration register allowing setting of various parameters in the system’s circuits and an output counter triggered by output pulses providing a full CT quantised 8-bit digital output.

### B. Design Considerations

Probably the most critical consideration of this design is ensuring that the amount of charge that is added to or removed from the balancing capacitor during a step up or down, \(\Delta Q^{\text{up}}\) and \(\Delta Q^{\text{down}}\) respectively is matched as much as possible.

Since both \(\Delta Q^{\text{up}}\) and \(\Delta Q^{\text{down}}\) are generated by a combination of a pulse of length \(t_p\) and injected current \(I_p\), there are two possible ways of ensuring their matching. Since it is generally complicated, if not impossible, to precisely match a current source and a current sink, a more feasible approach is varying the pulse length \(t_p\). This is achieved by a configurable 8-bit pulse length generator as outlined in Subsection III-E. Furthermore, additional flexibility in the size of \(\Delta Q^{\text{up}}\) and \(\Delta Q^{\text{down}}\) is provided by the use of a configurable current reference circuit described in Subsection III-G.

Even though this can lead to removal of a significant amount of mismatch, it is almost impossible to achieve absolutely equal \(\Delta Q^{\text{up}}\) and \(\Delta Q^{\text{down}}\). Their difference is however not necessarily very detrimental to the quality of the output signal as it only leads to addition of predictable distortion. It is entirely possible to calculate its magnitude in post-processing and completely remove it.

This is however to some extent made more complicated by the fact that additional unwanted additive component is present in the output due to leakage from capacitor \(C_b\). Even though this error is cumulative and increasing over time, it is also entirely predictable and not necessarily detrimental to the output SNDR. If we assume that the leakage current is independent from the input signal, this is only going to lead to addition of a slope \(k_l \times t\) to the output which can easily be estimated in a calibration process by sourcing a constant voltage to the input. As such leakage from capacitor \(C_b\) therefore only leads to a slight increase of sampling rate and hence an increase of idle power consumption due to introduction of a minimal achievable output pulse rate that is always present when quantising the leakage-incurred input signal.

Once \(k_l\) is established it is possible to estimate the mismatch \(\delta Q\) between \(\Delta Q^{\text{up}}\) and \(\Delta Q^{\text{down}}\) as

\[
\delta Q = \frac{\Delta Q^{\text{up}}}{\Delta Q^{\text{down}}} = \lim_{t \to \infty} \frac{N_{\text{up}} - k_l \times t}{N_{\text{down}}} \quad (5)
\]

where \(N_{\text{up}}\) and \(N_{\text{down}}\) are the amount of output spikes observed in a period of time. This paves a way to complete recovery of all detrimental effects due to capacitor leakage current and charge mismatch. A side effect of this added distortion presenting itself as a slope added to the output signal is a periodic overflowing of the output counter which has to be kept in mind and removed in post-processing.

Another critical part of the design is the input comparator. This is due to the fact that it typically has the highest power consumption of all the circuits in the system and determines the achieved performance in terms of bandwidth and resolution. It should be noted that generally CT-ADC implementations are not necessarily limited by bandwidth as is typical in discrete-time ADCs but rather by the slope or slew rate of the input signal.

This is due to the fact that there is only a certain amount of pulses that can be generated by the feedback loop of the system during a period of time. This constraint typically comes from the limited bandwidth of the input comparator.

The maximal input slew rate \(SR^{\text{max}}\) can be defined as

\[
SR^{\text{max}} = \frac{\Delta V_{\text{LSB}}}{t_s} \quad (6)
\]
where $t_s$ is the settling time of the feedback loop. To obtain an estimate of the needed comparator open-loop bandwidth $BW_c$, we can assume that the settling time of the comparator is roughly equal to three times the time constant $\tau$ of the amplifier and hence $t_s \approx 3\tau = 3/2\pi BW_c$.

To design the system for a certain constant 3-dB input bandwidth $B$, it is possible to do so by considering the absolute worst-case scenario - a pure tone signal $s(t)$ of frequency $B$ and amplitude $A/\sqrt{2}$ having a steepest slope of $ds/dt|_{\text{max}} = 2\pi BA/\sqrt{2}$. If we would like to achieve $N$-bit resolution the open-loop bandwidth of the comparator has to be approximately

$$BW_c \approx \frac{3}{2\pi t_s} = \frac{3SA_{\text{in}}^{\text{max}}}{2\pi V^{\text{LSB}}_{\text{in}}} = \frac{6\pi BA^N}{4\pi A\sqrt{2}}.$$  

(7)

This gives rise to fairly stringent demands on the amplifier input bandwidth as e.g. a 4 kHz, 8-bit implementation of the circuits requires an open-loop comparator bandwidth of $BW_c > 1$ MHz. It should however be noted that increasing the input frequency above the theoretical bandwidth limit of the circuit does not have the same effect as it would have in a conventional uniform sampling DAC where it would lead to aliasing. In this topology of CT ADC it would lead to distortion in the form of higher harmonics in the digitised signal very similar to the effect of finite slew-rate seen in operational amplifiers.

Even though not implemented in our design, the nature of (7) shows that this CT-ADC topology in theory has a capability of dynamically trading resolution for bandwidth. By altering the size of an LSB in line with Proposition I it is possible to decrease the resolution leading to a decrease of oversampling leading to no loss of information and an increase of maximal input bandwidth. This might be used in future versions of the circuit to implement an adaptive sampling scheme making use of Proposition I ensuring minimal oversampling improving the overall power efficiency.

C. Input Comparator

Since the presented implementation of the system is aimed at acquisition of neural signals (EAPs and LFPs), an analysis aimed at identifying the required bandwidth was carried out. A set of pre-recorded neural signals containing both LFPs and EAPs was scaled such that their peak-to-peak magnitude passed through 256 levels representing 8-bit quantisation. Differentiating those waveforms has shown a maximal slope not exceeding $2 \times 10^6$ steps/s leading to a requirement of $t_s < 500$ ns.

Using (7) it can then be shown that a comparator bandwidth of $BW_c \approx 1$ MHz is required to permit acquisition of this slope magnitude which is equivalent to designing for a worst-case input bandwidth of $B \approx 3.5$ kHz. To allow for a margin of safety, the system was designed for a maximal input 3 dB bandwidth of 4 kHz. It should be noted that even though this is smaller than maximal bandwidth of EAPs shown in Table I the performance of this design is sufficient as this topology of ADC is limited by the input slew-rate rather than input bandwidth which turns out to be advantageous in acquisition of neural signals.

A great challenge of the comparator implementation is coming from the fact that it is difficult to control its gain. It is almost impossible to use any feedback network as it is crucial to ensure that the inverting input has as high impedance as possible to ensure minimal charge leakage from $C_b$. In addition, the entire system loop has to have a gain at DC in order to ensure automatic set up of initial bias of $C_b$ during start up preventing the use of any capacitor-based feedback loops that might help overcoming the issue with leakage.

There is also a trade-off in the gain of the amplifier as too small a gain will lead to a poor noise performance of the circuit while a large gain will lead to a too high power consumption. This is especially the case as the system requires achieving a large open-loop bandwidth and therefore large gain would lead to a very large gain-bandwidth product.

As a result of these requirement our implementation uses a single-stage differential amplifier in an open-loop configuration leading to a gain of $A_c \approx 40$ dB and the required bandwidth at a bias current of 1.25 $\mu$A. The amplifier bandwidth is improved by using minimal size devices. This involves a trade-off between the overall static power consumption and amount of flicker noise. Using larger area of devices would linearly scale down flicker noise density which is desirable. Increasing device width however leads to an increase of parasitic output capacitance which has to be compensated by increase in bias current to preserve bandwidth. Similarly, increasing input pair length leads to higher output resistance, higher gain and hence smaller open-loop bandwidth. However, the input offset $V_{\text{off}}$, caused by possible poor mismatch between devices, is not of concern as it can be rejected by the loop of the DAC leading to the quantised voltage at $C_b$ being different by $V_{\text{off}}$ from the input.

The use of an open-loop amplifier leads to a variation of its gain due to temperature and manufacturing process variations. While the former is not of much concern due to the expected operation of the device in neural tissue and hence in stable temperature conditions, the latter is going to result in an introduction of a hysteresis between quantisation levels and its variation between different circuit samples.

The used threshold crossing detector, as described in the next subsection, indicates when the comparator output crosses one of two thresholds $V_{\text{SW}}^{\text{up}}$ or $V_{\text{SW}}^{\text{down}}$. The difference of those thresholds $\Delta V_{\text{SW}}^{c} = V_{\text{SW}}^{\text{up}} - V_{\text{SW}}^{\text{down}}$ referred to the input of the circuit,

$$\Delta V_{\text{SW}}^{\text{in}} = V_{\text{SW}}^{\text{up}} - V_{\text{SW}}^{\text{down}} / A_c,$$  

(8)

indicates a signal swing that has to be observed to trigger one level crossing up as well as down. Since this is not necessarily equal to one LSB, a hysteresis is inherently introduced.

This has a beneficial effect of preventing excessive switching in presence of noise but also a detrimental effect in the form of added distortion that can, however, be rectified in post-processing. It is possible to calculate the hysteresis magnitude and hence allow removal of the caused distortion from the quantised output by applying a suitable known input such as a sine wave during a calibration process. It however also has to be observed that the introduction of hysteresis leads to smaller overall sensitivity and in order to prevent aliasing it now has to be ensured that the highest-frequency content has peak-to-peak magnitude greater than $\Delta V_{\text{SW}}^{c}$ in line with Proposition I.
D. Threshold Crossing Detector

The output of the comparator is unity-gain buffered and connected to a threshold crossing detector indicating whenever one of the quantisation levels is crossed. As shown in Fig. 9 this is composed of two inverters making use of different transistor lengths controlling their switching points $V_{\text{up}}$ and $V_{\text{down}}$. If the difference $\Delta V_{\text{sw}} = V_{\text{up}} - V_{\text{down}}$ is set such that $\Delta V_{\text{sw}} = A_c \times \Delta V_{\text{LSB}}$ the circuit implements a CT digitizer with fixed quantisation levels. As discussed earlier, it is however practically impossible to ensure that this is satisfied due to dependence on several design parameters prone to process, voltage and temperature (PVT) variations leading to a difference of $\Delta V_{\text{sw}} - A_c \times \Delta V_{\text{LSB}}$ introducing a hysteresis.

Since the hysteresis magnitude affects the sensitivity of the circuit and can be important for its operation, its control is enabled by electronically altering the length of crossing detector transistors using configuration stored in a shift register. Each $V_{\text{sw}}$ is controlled by three thermometer-coded bits allowing four different settings.

E. Variable Pulse Generator

As outlined earlier, a critical part of the design is a variable pulse length generator allowing control of mismatch between added and removed charge $\Delta Q_{\text{up}}$ and $\Delta Q_{\text{down}}$. As seen in Fig. 10 this is composed of a chain of inverter-based delay lines and multiplexers allowing configuration of both the charging and discharging pulse length in the region of $\approx 50 \text{ ns} < t_p < 110 \text{ ns}$ permitting rectification of current source or sink mismatch of up to 120% found by Monte Carlo simulations.

The resolution of each pulse length configuration is 8 bits. The weighting of bits has however been chosen as 1.7 rather than binary 2 in order to prevent gaps in possible configuration levels due to process variations leading to delay mismatch between subsequent delay elements. The used weighting can however lead to non-linearity of the $t_p$ setting curve which must be taken into consideration when designing a suitable calibration algorithm.

In addition, the pulse generator is complemented by an inhibitor circuit introducing a 2-bit configuration of a dead time $t_d$ that prevents generation of two subsequent pulses in a period shorter than $\approx 163 \text{ ns} < t_d < 632 \text{ ns}$. This ensures stability of the system as the response of the input comparator could be slower than $t_p$ leading to multiple pulses being generated after each quantisation level crossing.

F. Precise Charge Injection

Another challenge lies in ensuring that the amount of extra injected charge $\delta Q$ due to the operation of the current source and sink is as small as possible. This is achieved by circuit seen in Fig. 11 ensuring uninterrupted flow of current and bootstrapping the drain of the current source to the same potential as $C_b$ preventing charge build-up on the switch. Additional dummy transistors (not shown in Fig. 11) as well as use of small devices reduce the effect of charge injection from the channel of the used active switches. Miller compensation is used to ensure stability of the bootstrapping loop.

This reduces the extra injected charge to the order of hundreds of $aC$ which can be rectified by a slight modification of pulse length $t_p$ whose flexibility is sufficient to remove both mismatch between source and sink currents as well as errors caused by extra charge injection.

G. Current Reference

A resistor-less self-biasing current reference shown in [42] was used for generation of reference current used for biasing of all circuits as well as for generation of charging and discharging currents $I_{\text{up}}$ and $I_{\text{down}}$ used in creation of charge packets $\Delta Q_{\text{up}}$ and $\Delta Q_{\text{down}}$. Those currents are both configurable in the range...
of 4–32 nA in steps of 4 nA. Since Monte-Carlo simulations have shown that the variable pulse-length generators alone are sufficient for removing charge mismatch the reference circuit configuration was designed to only allow configurations where $I_{up} = I_{down}$.

Since $C_b$ was chosen as $C_b = 3 \, \text{pF}$ the size of one LSB is configurable in the range between $V_{LSB}^{\text{min}}$ and $V_{LSB}^{\text{max}}$ given as:

$$V_{LSB}^{\text{min}} = \frac{I_{p}^{\text{min}} \times t_{p}^{\text{min}}}{C_b} = \frac{4 \, \text{nA} \times 50 \, \text{ns}}{3 \, \text{pF}} = 66.6 \, \mu\text{V}$$  \hspace{1cm} (9)

$$V_{LSB}^{\text{max}} = \frac{I_{p}^{\text{max}} \times t_{p}^{\text{max}}}{C_b} = \frac{32 \, \text{nA} \times 110 \, \text{ns}}{3 \, \text{pF}} = 1.17 \, \text{mV}$$  \hspace{1cm} (10)

### IV. Measurement Results

The described system was implemented in a commercially available 0.35 $\mu$m CMOS technology operating from a supply voltage of 1.5 V. The core occupying an area of $\approx 320 \, \mu\text{m} \times 360 \, \mu\text{m}$ is illustrated in Fig. 12 alongside a microphotograph of the fabricated die.

The static power consumption was measured as $\approx 3.75 \, \mu\text{W}$. To verify operation of the circuit the system was configured such that $I_{p}^{\text{down}} = I_{p}^{\text{up}} = 16 \, \text{nA}$, maximal possible $\Delta V_{sw}$ and shortest possible $t_p$. A sinusoidal input passing through $\approx 256$ quantisation levels ($\approx 8$ bits) was then applied to the input and the output in form of spikes recorded at a sampling rate of 100 MS/s using a Salae Logic Pro 8 logic analyser.

The output was reconstructed using spline interpolation and mismatch removal based on (5). The frequency of the input signal was then varied between 70 Hz and 4036 Hz and average power consumption and SNDR recorded as seen in Fig. 13. This shows linear dependence of power consumption on the input frequency demonstrating the anticipated activity-dependent operation.

The SNDR has a peak at input frequencies around 1.7 kHz and drops for higher and smaller frequencies. It should however be noted that the recorded SNDR is showing a situation when the input is a single tone and as such the SNDR profile is dependent on the input signal. The output power spectra plotted in Figs. 14 and 15 showing an input of 1794 Hz and 531 Hz respectively demonstrates a difference in noise profile as well as higher distortion when digitising signals containing only small frequencies. Measurements and additional simulations have shown that this is caused by ringing observed in the charge injection bootstrapping feedback loop shown in Fig. 11. This only presents itself during periods of small input signal slope and the corresponding small output pulse rate. Since the instability is input signal dependent, it leads to increased distortion and raised noise floor as seen in Fig. 15. This is caused by the fact that this results in introduction of sharp semi-random peaks in the signal as seen in Fig. 16. As this phenomena only presents itself when a signal of small input slope is digitised, a higher noise floor is observed during acquisition of a 531 Hz sine wave when compared to 1794 Hz.

A peak SFDR of 38.8 dBc and SNDR (measured between 100 Hz and 10 kHz) of 28.7 dB was observed at an input frequency of 1794 Hz. This is equivalent to a maximal observed ENOB of 4.47 bits. To investigate the effect of current leakage...
<table>
<thead>
<tr>
<th></th>
<th>[units]</th>
<th>[14]</th>
<th>[18]</th>
<th>[15]</th>
<th>[16]</th>
<th>[17]</th>
<th>[19]</th>
<th>[20]</th>
<th>[21]</th>
<th>[22]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>N [bits]</td>
<td>8</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>4–8</td>
<td>5</td>
<td>–</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>SFDR [dB]</td>
<td>58.3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>22–38.8</td>
<td></td>
</tr>
<tr>
<td>ENOB [bits]</td>
<td>7.5–8.69</td>
<td>4.42</td>
<td>6.36–7.86</td>
<td>8.26</td>
<td>7.35–8.02</td>
<td>5.84–7.69</td>
<td>6.89</td>
<td>4.87</td>
<td>5–6.7</td>
<td>2.3–4.8</td>
<td></td>
</tr>
<tr>
<td>BW [kHz]</td>
<td>20</td>
<td>≥0.35</td>
<td>5</td>
<td>0.11–10.5</td>
<td>0.95</td>
<td>1</td>
<td>1</td>
<td>0.2–5</td>
<td>10^4–5 × 10^4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Tech [μm]</td>
<td>0.13</td>
<td>0.13</td>
<td>0.18</td>
<td>0.09</td>
<td>0.18</td>
<td>0.35</td>
<td>0.18</td>
<td>0.5</td>
<td>0.028</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Area [mm^2]</td>
<td>1.69</td>
<td>0.357</td>
<td>0.045</td>
<td>N/A</td>
<td>0.49</td>
<td>0.037</td>
<td>0.96</td>
<td>0.06</td>
<td>0.0032</td>
<td>0.115</td>
<td></td>
</tr>
<tr>
<td>Supply [V]</td>
<td>0.8</td>
<td>0.3</td>
<td>0.8</td>
<td>0.5 &amp; 0.7</td>
<td>1</td>
<td>1.8–2.4</td>
<td>0.7</td>
<td>3.3</td>
<td>0.65</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Power [μW]</td>
<td>3–9</td>
<td>0.22</td>
<td>0.31–0.58</td>
<td>0.54–0.73</td>
<td>8.49</td>
<td>0.6–2.0</td>
<td>25</td>
<td>118.8–501.6</td>
<td>8–24</td>
<td>≈3.75–12.9</td>
<td></td>
</tr>
<tr>
<td>FOM [pJ]</td>
<td>0.36–2.49</td>
<td>29.36</td>
<td>0.266–1.41</td>
<td>0.168–21.64</td>
<td>34.9–54.77</td>
<td>2.9–34.9</td>
<td>210.8</td>
<td>812.5–85778</td>
<td>(1.5–75) × 10^{-3}</td>
<td>33.7–654.9</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 16. Reconstructed 200 mV, 531 Hz sine wave passing through 256 quantisation levels (8 bits) showing distortion and noise present during small-slope intervals of the input signal.

Fig. 17. An example of a reconstructed neural spike (dashed) and the original signal (solid) passing through 164 quantisation levels.

from capacitor $C_b$, a fixed voltage was applied to the input. This resulted in a pulse indicating a shift up being generated approximately once every second. This shows that the leakage has an effect of adding a slope of no more than a couple of LSB per second to the output signal which in most applications is negligible. A sample neural signal (EAP + LFP) amplified such that it passes ≈256 quantisation levels was then applied as an input to observe behaviour of the system in a more realistic scenario. While the non-ideality of the circuit leads to significant distortion during quiet periods of the signal containing only LFPs, the circuit showed an ability to successfully digitise neural spikes that contain steep slopes as demonstrated in Fig. 17 showing a reconstructed neural spike passing through 164 quantisation levels leading to a mean squared error (MSE) of 3.5 bits. The SNDR as defined by (3) was found to be 30.62 dB (equivalent to ENOB of 4.8 bits) in this case.

In addition, it was verified that the minimal period between passing quantisation levels is ≈360 ns which as demonstrated in Subsection III-C is sufficient for 8-bit acquisition of EAP signals. Acquisition of a typical neural signal led to an average power consumption of ≈4.2 μW showing a slight increase form the static value of ≈3.75 μW demonstrating the efficiency of the system in digitisation of neural signals.

V. CONCLUSION

This work has demonstrated that using non-uniform sampling can lead to significant savings during digitisation of certain sparse biological signals and compression ratios exceeding 20 can be achieved during acquisition of EAP signals.

To compare the presented system with alternative implementations of CT-ADC we can use a Figure of Merit (FOM) defined as:

$$ FOM = \frac{P}{2^{ENOB} \times BW} $$

where $P$ is the static power consumption, $ENOB$ the effective amount of bits and $BW$ the designed bandwidth. A design achieving a smaller FOM is comparably better.

The system presented herein is compared to state-of-the-art CT-ADC implementations that have been reported in the literature, in Table II. To the best knowledge of the authors it currently shows the best FOM of a truly CT system implemented in a 0.35 μm CMOS technology. Even though the results reported in [19] lead to a slightly superior FOM, the presented system uses a clocked comparator to achieve a smaller power consumption which means that the system is not truly continuous-time in nature.
The achieved FOM of the presented system is significantly affected by the observed flaw of the circuit causing ringing in the charge injection circuit and leading to a reduced ENOB. This could be easily rectified without causing an additional increase of power consumption by increasing the size of the used compensation capacitor. As seen in Fig. 18, the majority of the observed power consumption (2.65 μW) is attributed to the input comparator and its unity-gain buffer. Without modifying the circuit, this would significantly benefit of a finer CMOS technology with smaller threshold voltages allowing a reduction of supply voltage.

It should also be noted that the reported power consumption of our system is inflated by ≈1.5 μW due to inclusion of additional buffers that are used exclusively for debugging and testing purposes. The system presented herein provides a proof of concept of the proposed topology and has areas of possible improvement to be done in the future.

Those could improve the static power consumption by utilizing a more sophisticated comparator such as one employing dynamic biasing. Furthermore, dynamic power consumption can be significantly improved by implementing in a finer pitch technology – as with the exception of the input comparator the system is completely digital in nature.

A competitive advantage of the presented topology is its compact area. This is expected to scale with the used feature size as a large majority of the used circuits are digital. In addition, the manufactured layout seen in Fig. 12 could be further optimised to occupy a smaller area without changing the used technology.

Since the presented circuit operates with input-referred LSB size in the range of hundreds of μV it would have to be complemented by a suitable neural amplifier to form a complete neural acquisition system. An interesting complement might be a folding amplifier presented in [43] exploiting 1/f characteristics of neural signals. The folding nature would allow to relax the linearity requirements on the input ADC comparator while adding the possibility of combining digital output signals of both these circuits to increase the overall dynamic range.

ACKNOWLEDGMENT

The authors would like to thank Dr. L. Leene for helpful comments and assistance with improving this manuscript.

REFERENCES


Michael Maslik (S’17) received the M.Eng. degree in electrical and electronic engineering from Imperial College London, London, U.K., in 2016, where he is currently working toward the Ph.D. degree. His research interests include development of microelectronics for use in distributed millimeter-scale neural implants, development of instrumentation circuits and compression methods allowing acquisition of local field potentials and their use in brain machine interfaces.

Yan Liu (M’08–M’12) received the B.Eng. and M.Sc. degrees from Zhejiang University, Hangzhou, China, in 2006 and 2007, respectively, and the Ph.D. degree in electrical and electronic engineering from Imperial College London, London, U.K., in 2012. He is a Research Fellow with the Centre of Bio-Inspired Technology, Electrical and Electronic Engineering, Imperial College London. His research interests include CMOS-based lab-on-chip devices and platforms, brain machine interface, and novel mixed signal circuits for biomedical applications.

Tor Sverre (Bassan) Lande (M’93–SM’06–F’10) is a Professor of microelectronics at the Department of Informatics, University of Oslo, Oslo, Norway, as well as a Visiting Professor at the Institute of Biomedical Engineering, Imperial College London, London, U.K. He is the author or coauthor of more than 150 scientific publications with chapters in three books. His research interests include microelectronics, both digital and analog, neuroengineering, analog signal processing, subthreshold circuit and system design, biomedical circuits and systems, RF CMOS, pulsed radio (UWB), and radar systems. He is an Associate Editor of several scientific journals and as a Guest Editor for several IEEE journals. He is a Technical Committee Member of several international conferences and has served as a Reviewer for a number of international technical journals. He has served as the Technical Program Chair for several international conferences (ISCAS 2003 in Bangkok, NORCHIP, BioCAS2004, BIOCAS2006, BioCAS2010, ISCAS 2011, and ISCAS 2017). He was the Chair Elect (2003–2005) of the IEEE Biomedical Circuits and Systems technical committee (BioCAS) and is also a Member of other CAS technical committees. In 2006, he was appointed as the Distinguished Lecturer of the IEEE Circuits and Systems Society (CAS) and an Elected Member of CAS Board of Governors. He has served as the CAS Vice-President of conferences (2011–2013). He is also the founding Editor-in-Chief of IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS (2007–2010). He is a Member of the Norwegian Academy of Technical Sciences.

Timothy G. Constandinou (AM’98–M’01–SM’10) received the B.Eng. and Ph.D. degrees in electronic engineering from Imperial College London, in 2001 and 2005, respectively. He is currently a Reader of neural microsystems within the Circuits and Systems Group, Department of Electrical and Electronic Engineering, Imperial College London and also the Deputy Director of the Centre for Bio-Inspired Technology. His research interests include neural microsystems, neural prosthetics, brain-machine interfaces, implantable devices, and low-power microelectronics. Within the IEEE, he serves on several committees/panels, regularly contributing to conference organization, technical activities, and governance. He is the Chair of the IEEE Sensory Systems Technical Committee, is a Member of the IEEE BioCAS Technical Committee, IEEE Brain Initiative Steering Committee, and serves on the IEEE Circuits and Systems Society Board of Governors for the term 2017–2019. He was the Technical Program Chair of the 2010, 2011, and 2018 IEEE BioCAS conferences, the General Chair of the BrainCAS 2016 and NeuroCAS 2018 workshops, the Special Session Co-Chair of the 2017 IEEE ISCAS Conference, and the Demonstrations Co-Chair of the 2017 BioCAS Conference. He is currently an Associate Editor of the IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS. He is a Fellow of the IET, a Chartered Engineer, and Member of the IoP.