

Analog Neuromorphic Systems for Implantable Neural Interfaces

Arindam Basu

arinbasu@cityu.edu.hk

City University of Hong Kong <https://orcid.org/0000-0003-1035-8770>

Pao-Sheng Vincent Sun

City University of Hong Kong

Ye Ke

City University of Hong Kong

Zhengnan Fu

City University of Hong Kong

Elias Arnold

Heidelberg University

Chiara Bartolozzi

Istituto Italiano di Tecnologia <https://orcid.org/0000-0003-3465-6449>

Yu-Lin Chang

National Taiwan University of Science and Technology

Chiara De Luca

Digital Society Initiative, University of Zurich <https://orcid.org/0000-0003-3488-0088>

Elisa Donati

University of Zurich and ETHZ <https://orcid.org/0000-0002-8091-1298>

Olympia Gallou

Institute of Neuroinformatics, University of Zurich and ETH Zurich

Ashish Gautam

Computer Science and Mathematics Division, Oak Ridge National Laboratory

Quentin Halbach

Event-Driven Perception in Robotics, Istituto Italiano di Tecnologia

Chun-Yen Huang

National Yang Ming Chiao Tung University

Chieh-Hsi Kuo

National Taiwan University of Science and Technology

Kaushik Lakshmiramanan

University of Maryland

Saptarshi Maiti

Indian Institute of Science <https://orcid.org/0000-0002-6557-8626>

Eric Müller

Heidelberg University <https://orcid.org/0000-0001-5880-2012>

Sheng-Yu Peng

National Yang Ming Chiao Tung University

Shavika Rastogi

Western Sydney University

Chetan Thakur

Indian Institute of Science <https://orcid.org/0000-0002-1240-6214>

Anubhab Tripathi

City University of Hong Kong

Abhinav Uppal

University of California, San Diego <https://orcid.org/0009-0006-4722-1005>

Biyan Zhou

City University of Hong Kong

Jennifer Hasler

Georgia Institute of Technology

Giacomo Indiveri

University of Zurich and ETH Zurich <https://orcid.org/0000-0002-7109-1689>

Johannes Schemmel

Heidelberg University

Saptarshi Ghosh

Institute of Neuroinformatics, University of Zurich and ETH Zurich <https://orcid.org/0000-0001-8861-5269>

Article

Keywords:

Posted Date: September 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10913660/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Analog Neuromorphic Systems for Implantable Neural Interfaces

Pao-Sheng Vincent Sun^{1,*}, Ye Ke^{1,*}, Zhengnan Fu¹, Elias Arnold⁹, Chiara Bartolozzi⁴, Yu-Lin Chang¹³, Chiara De Luca^{2,12}, Elisa Donati², Olympia Gallou², Ashish Gautam⁸, Quentin Halbach⁴, Chun-Yen Huang⁵, Chieh-Hsi Kuo¹³, Kaushik Lakshmiramanan¹⁰, Saptarshi Maiti⁷, Eric Mueller⁹, Sheng-Yu Peng⁵, Shavika Rastogi¹⁴, Chetan Singh Thakur⁷, Anubhab Tripathi¹, Abhinav Uppal¹¹, Biyan Zhou¹, Jennifer Hasler⁶, Giacomo Indiveri², Johannes Schemmel⁹, Saptarshi Ghosh^{2,+}, and Arindam Basu^{1,3+}

¹Department of Electrical Engineering, City University of Hong Kong, Hong Kong SAR, China.

²Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland

³State Key Laboratory of Terahertz and Millimeter Waves, City University of Hong Kong, Hong Kong SAR, China.

⁴Event-Driven Perception in Robotics, Istituto Italiano di Tecnologia, Genoa, Italy

⁵National Yang Ming Chiao Tung University, Taiwan

⁶School of ECE, Georgia Institute of Technology, USA.

⁷Department of Electronic Systems Engineering, Indian Institute of Science, India

⁸Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, USA.

⁹Heidelberg University, Germany

¹⁰Department of Electrical and Computer Engineering, University of Maryland, College Park, USA

¹¹Shu Chien - Gene Lay Department of Bioengineering, University of California San Diego, La Jolla, CA, USA.

¹²Digital Society Initiative, University of Zurich, Zurich, Switzerland.

¹³National Taiwan University of Science and Technology, Taiwan.

¹⁴Western Sydney University, Australia

* Equal Contribution

+Correspondence to arinbasu@cityu.edu.hk and sapta@ini.uzh.ch

ABSTRACT

With the increasing popularity of brain-machine interfaces for a variety of motor disorders and fundamental neuroscientific experiments, there is also a growth in new sensors with increasing number of parallel channels to sense brain activity. Hence, there is an urgent need to manage the exploding data rate and power dissipation of these devices. Analog neuromorphic computing promises to be a natural candidate for bringing intelligence to implantable neural interfaces, thus only needing to transmit sparse information gleaned from the raw neural recordings. Neuromorphic event-driven sampling can reduce data rate at the source producing spikes that can natively be processed by low-power, sparsely activated spiking neural networks (SNN), that detect events of interest such as epileptic seizures or occurrence of biological action potentials (AP). This work compares popular neuromorphic analog-spike converters (ASC) in terms of their robustness to interference and shows delta modulation to be superior to integrate and fire modes of spike generation. Further, a new type of delta modulation inspired by high threshold bursting neurons in the hippocampus is shown to extend the tradeoff between compression ratio and data fidelity for conventional delta modulation. Measured results from fabricated integrated circuits show excellent data compression rates far exceeding those of conventional analog digital converter with Nyquist sampling. Further, mixed-signal, clockless SNNs interfaced with this bursting ASC are shown to be able to classify the generated spikes to detect AP events with high accuracy while dissipating scant energy.

Introduction

The rapid growth of Artificial Intelligence (AI) algorithms, electronic processors and development of new sensors and improvements in connectivity have converged to create an explosion in the Artificial Intelligence of Things (AIoT). Along with a rise in smart machines such as autonomous vehicles, intelligent robots or smart surveillance systems, there has been a steady rise in battery operated, body-worn smart electronics for a variety of applications. On one hand we have smart glasses, AR/VR headsets and smart watches being used ubiquitously for consumer applications, while on the other hand we have a variety of biosensors that are being used to monitor our body state in medical electronics through different signals such as

electroencephalogram (EEG)¹ or neural action potentials (AP)² for brain monitoring, electrocardiogram (ECG) for cardiac signals³ as well as glucose⁴ and insulin for diabetic patients. In fact biomedical signals such as electromyography (EMG) have also been used for human-computer interaction⁵. A lot of these devices, such as image sensors for smart glasses and microelectrode arrays (MEA) for brain monitoring face a data deluge problem due to an increasing number of pixels requiring innovations in compression and communication schemes.

One prominent example of multi-channel brain sensors is the implantable microelectrode arrays used in communication or movement Brain-Machine Interfaces (BMI)⁶, where the sensor can, for example, be used to capture neural activity in motor cortex of a tetraplegic patient, which can be decoded to a command for a machine such as a computer, wheelchair or robotic arm. Implanted BMIs for communication were previously used to spell words by selecting characters on a computer screen using a brain-controlled cursor⁷⁻⁹ and provided much higher decoded character rates than their wearable EEG speller counterparts. Recent advances¹⁰⁻¹² in direct decoding of attempted speech have increased this word rate close to conversational speech. However, in all these cases, having a wired connection from the implant increases risk of infection and patient discomfort. Hence, wireless implants are necessary for commercial translation and has been the focus of much research in the past few years.

Another important usage of a multi-channel brain sensor is in therapeutic BCIs to record intracranial EEG (iEEG) signals, from subdural grids and strips on the cortical surface (ECoG) or from stereotactically placed depth electrodes (sEEG), to locate the epileptogenic zone for patients with pharmaco-resistant focal epilepsy who might undergo resective surgery. Locating the seizure onset zone(s) can inform surgeons about potential resection areas¹³; however, because seizures are rare and unpredictable, locating this zone requires continuous recordings over one or more weeks, during which the patient remains tethered to a recording system in a monitoring unit in a clinical facility. The demand is also asymmetric, since tens or hundreds of contacts must be digitized continuously at kilohertz sampling rates for days, while the clinically meaningful events (seizures) occupy only a negligible fraction of the record. Chronic ambulatory iEEG is already possible via implanted neurostimulators and, less invasively, via subscalp systems. However, such systems, currently in clinical use, can only record a handful of channels and store intermittent snippets rather than continuous data¹⁴. Therefore, an wireless and compact device would free these patients from prolonged in-hospital settings and allow the continuous long-term collection of ambulatory iEEG data.

Creating wireless implants is challenging due to the tight area and power constraints of implanted electronics to limit significant tissue heating and frequent battery recharging. The high data rate further requires high transmission power to reduce bit error rates. The problem is further exacerbated by the growing number of channels in implantable neural probes (e.g. 384 channels to record from 5120 physical sites in Neuropixels 2.0 NHP¹⁵), that support a Moore's law like exponential increase in the number of simultaneously recorded neurons¹⁶. A similar trend exists for ECoG measurements with recent work extending to 1024 to 2048 electrodes¹⁷. While new wireless transmission methods are being developed^{18,19} to support high bandwidth data transmission, there is a strong need to develop low-power, low-area in-sensor data compression and processing circuits to reduce the datarate at the source, while respecting the strict resource constraints of implantable electronics.

Interestingly, the biological nervous system provides a great example of handling data efficiently from millions of sensors (e.g. photoreceptors in the retina or mechanoreceptors in the skin) where the sensors provide in-sensor adaptation and compression by change detection²⁰. The cortical neurons further recognize events of interest using very low-power analog networks of spiking neurons with rich local dynamics conferred by a multitude of ion-channels and robust communication using digital-like action potentials or spikes. Inspired by this, neuromorphic sensors and processors have emerged for low-latency, data and power efficient sensing and processing^{21,22}. Event-based sensors^{23,24} respond to changes in signal, and only emit asynchronous events signaling the time and polarity of such changes. This compressed representations enables sparse communication with neuromorphic processors which can use analog^{25,26} or digital²⁷ neurons, while communicating between neurons using similar asynchronous event-based communication principles. Thus, neuromorphic sensing and processing opens up a new avenue for efficient and scalable systems for brain activity monitoring.

Conventional implanted neural interfaces typically comprise a bandpass amplifier(s) with variable gain followed by multiplexed analog-to-digital converters (ADC) that operate at a Nyquist rate, set by the fastest possible signal to be encountered. However, events of interest, such as action potentials (AP) happen rarely due to the sparsity of neural activity, resulting in a lot of energy and bandwidth wasted in digitizing irrelevant signals. In contrast, neuromorphic neural interfaces (NNI) that convert analog signals to digital spikes only when "interesting" events occur, could provide data savings proportional to event sparsity. As an inspiration, the human retina uses retinal ganglion cells to send spikes for downstream processing only when the input from the bipolar cells exceeds the spatio-temporal average by a large enough amount, with the sign of the change being conveyed by two separate (ON and OFF) classes of cells. Inspired by this, NNI for interfacing with high density MEAs for in-vitro studies²⁸ have used Delta modulators (DM) with dual polarity outputs for data efficient digitization (Fig. 1c). Other efforts^{29,30} has used similar asynchronous DM circuits to digitize ECoG signals with the goal of efficiently detect seizures and high frequency oscillations (HFO) in an attempt to locate epileptogenic zones. A different class of analog to spike converters (ASC) takes inspiration from cortical neurons, and capture their simplified dynamics (in proximity of a saddle node bifurcation)

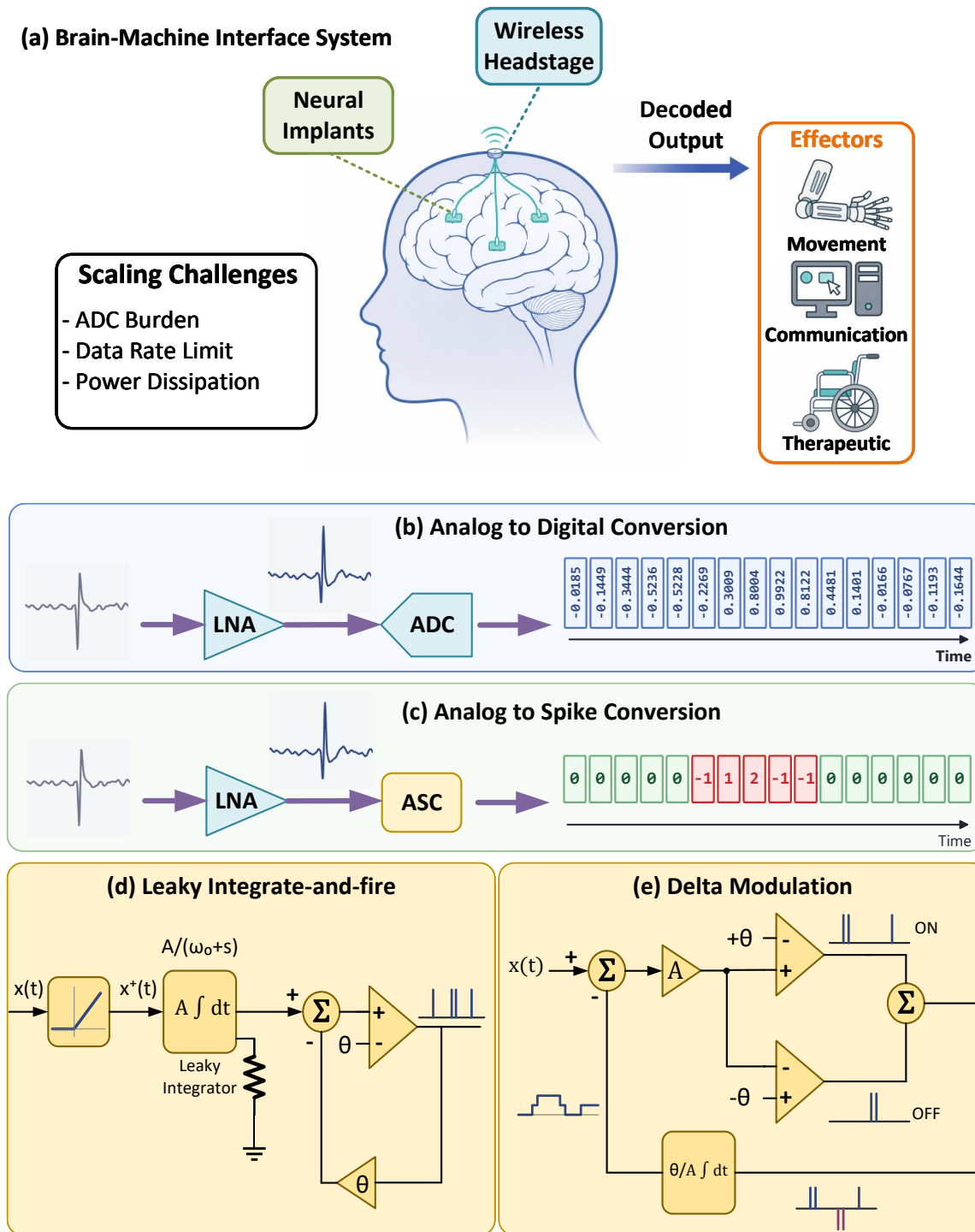


Fig. 1. **a** A conventional BMI system comprises an implant that communicates with effectors (typically via a headstage) such as computers, wheelchairs or robotic limbs. Converting the links to wireless ones for patient comfort is faced with power-dissipation challenges related to data digitization and transmission. **b** Conventional Analog-to-digital conversion (ADC) samples the neural signal at regular times leading to a data deluge. **c** In contrast, neuromorphic sampling or Analog-to-spike conversion (ASC) is capable of reducing the data rate by sampling only when the signal changes beyond a threshold, leading to signal sparsity driven compression at the source. In this case, summed event counts in a time bin are shown. Block diagram of two typically used neuromorphic ASC: **d** a leaky integrate-and-fire (LIF) neuron and **e** a delta modulator (DM), both of which convert analog signals to a sequence of sparse events. The LIF requires another pipeline similar to the one shown for handling the negative excursions of the input analog signal. Parts of this figure were enhanced using AI.

using a first order differential equation mimicking low-pass filtering, followed by a threshold to generate digital spikes (Fig. 1d). Referred to as the leaky integrate and fire (LIF) model, it has been used in some works^{31,32} for generating spike encoded temporal features for ECG and EMG signals. However, there has not been a principled comparison between these popular ASC schemes used in NNIs in terms of their robustness to interference, data compression and information content for downstream tasks. Foundational work^{29–31} for single or limited channels have been presented; however, the scalability of their approach to future generation of high density brain sensing remains unexplored. The approach of using cochlea inspired ASC using LIF³² requires additional bank of bandpass filters in each channel that increases power and area significantly for high density probes. While an approach²⁸ has demonstrated a high density NNI for in-vitro sensing, it uses a conventional DM approach that suffers from strict tradeoff between noise induced spurious spikes and accurate signal reconstruction. Moreover, the single ended circuits used in this work will be severely affected by the interference and common-mode noise observed in implanted electronics. Advanced DM circuits with adaptive steps such as Adaptive Asynchronous Delta Modulation (aADM)^{33,34} or pixels with contrastive pre-processing^{35–37} have been reported in literature to mitigate the shortcomings of classical DM. However, they typically entail significant hardware overhead and increase the area per channel, presenting scaling challenges.

In this work, we first show that the DM approach is inherently more robust than LIF to large magnitudes of interfering signals when designed to meet the strict area constraints of multi-channel, high density neural interfaces. We demonstrate the possibility to extract neural AP and Local field potentials (LFP) from digital spikes generated by a single bandpass DM circuit, whereas LIF ASC modules require separate filters for both neural signals. The ability to capture both spikes and LFP are crucial for high-performance neural interfaces due to the complementary information contained in both³⁸, higher robustness of LFP based decoders and recent improvements in decoding accuracy from these decoders³⁹. We also show > 10 dB improvement in rejecting power-line interference for DM-based encoders. However, DM-based ASC are known to have a strict tradeoff between accuracy and compression according to rate distortion theory. Hence, we introduce a novel ASC for NNIs inspired by bursting neurons in mammalian cortex. In bursting neurons, the fast spiking activity is modulated by a slow variable such as activation of an outward current (e.g. due to persistent K^+ channels) or inactivation of an inward current (e.g. transient Ca^{2+} channels)⁴⁰. This slow variable causes the cell to generate a burst and then transition to silence. Similarly, we developed a Burst Delta Modulation (BDM) circuit where the ASC generates spikes only if a large threshold is crossed. The duration of bursts is governed by a “slow” process emulated by digital circuits which again transitions the ASC to silence. To evaluate the efficacy of this approach, we go beyond traditional reconstruction errors and use two very different detection tasks: detecting neural AP in noisy intracortical recordings and detecting epileptic seizures from iEEG signals. The tasks were chosen to represent two extremes of neural dynamics—AP are high-frequency, short and sparse transient events of interest, while seizures have more mixed dynamics with longer, lower-frequency state transitions. We demonstrate $> 1000X$ compression for intracortical AP detection and $\approx 2000X$ compression for seizure detection for these ASC over conventional ADC. Finally, we validated our results using a two chip neuromorphic system: a 65 nm CMOS design of the BDM that achieves energy efficiency of $2.85 \mu W/\text{channel}$ and an accelerated analog spiking neural network chip in 65 nm CMOS with energy efficiency of ≈ 10 pJ/spike. Further, we show BDM provides $> 5.5X$ compression at similar RMSE compared to conventional DM. Evaluated on the AP detection task, the combined system achieves excellent accuracy of 93 – 96% and compression of 13 – 37x for SNR levels ranging from 14.86 – 16.43 dB while the acceleration factor enables sharing among > 20 channels. Compared with conventional ADC-based neural interfaces, the neuromorphic neural interface provides measured compression rates of 4130x after the ASC, which is 40 – 300x better than prior art, providing a pathway for high-density, wireless neural interfaces of the future.

Results

Robustness to Local Field Potentials and Power-line interference

Two types of neural signals are commonly used for decoding and analysis—low-bandwidth ($\approx 1 - 100$ Hz) local field potentials (LFP) indicative of population neural activity and higher bandwidth ($\approx 300 - 5000$ Hz) action potentials (AP) reflective of single neuron excitability. Both signals have been shown to carry different information⁴¹, and even combining the two modes by considering AP firing times with respect to the phase of the underlying LFP was found to be more informative than either mode alone³⁸. Conventionally, this was handled by having separate filters for LFP and AP in each channel of the neural interface. However, given the increasing number of channels in modern neural probes¹⁵, there is a trend of decreasing area per channel in modern large scale neural interfaces (Fig. 2a). Hence, we propose an architecture of the NNI where each channel has a single broadband LNA covering both AP and LFP bands, followed by a single ASC with the potential of extracting AP and LFP by processing the combined events generated by the ASC.

In this context, we evaluate the two most popularly used ASC techniques—Leaky Integrate and Fire (LIF), which encodes the absolute amplitude of the signal and Delta Modulator (DM), which encodes the changes in the amplitude of the signal. Fig. 1d and e show the mechanism of spike generation in these two techniques. DM^{42,43} is an ASC involving a level shifting approach that outputs a spike or event based on the difference in the amplitude of a signal. If a signal shows a change (denoted

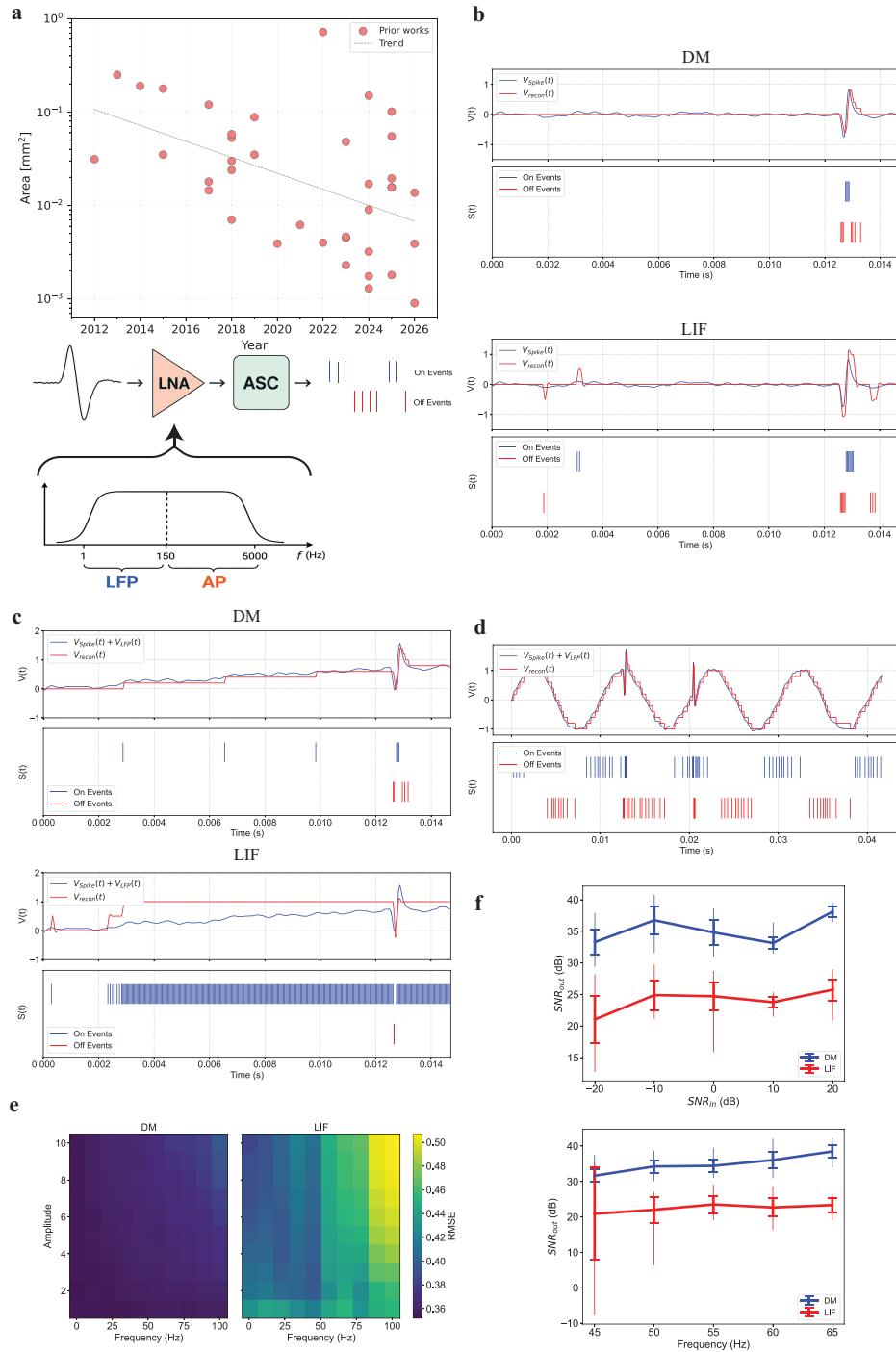


Fig. 2. **a** The trend for neural interface circuits published in the last decade shows a clear reduction in area per pixel as we move towards higher channel counts. Hence, a single pipeline of LNA and ASC for both LFP and AP recording are preferred to reduce area. **b** Comparison of DM and LIF encoding on a snippet of neural recording (with the LFP filtered out) showing the generation of events of dual polarity. The reconstructed waveform overlayed in red shows high fidelity. **c** Comparison of DM and LIF encoding for the same snippet but with an added 10 Hz sine wave to mimic low frequency LFP or interference signals such as motion artifacts. While the output from DM can still reconstruct the AP, LIF is overwhelmed by the interference. **d** On a longer time scale, DM is capable of capturing both LFP and AP in neural signals, with highly accurate reconstruction. **e** Effect of wide variety of low frequency (1-100 Hz) noise with up to ten times AP amplitude on DM and LIF encoding exhibiting consistently lower AP reconstruction RMSE for DM compared to LIF. **f** The SNR of DM and LIF encoding of varying quality of ECoG signals (SNR varying from -20 to 20 dB) tempered with power line interference of 60 Hz can be improved after reconstruction by dedicated algorithms, with DM showing consistently >10 dB improvement over LIF. The same conclusion holds true for different power line frequencies.

as delta Δ) in amplitude that is more than a preset threshold (denoted as theta θ), the encoder produces a spike/event. An ‘On’ or ‘Off’ label is attached based on the polarity of the change, i.e. if the signal has increased or decreased at the level crossing, respectively. Hence, DM can naturally handle bipolar analog inputs by producing events with dual polarities, indicating up- and down-ward changes, respectively.

LIF⁴⁴ is an ASC approach where a signal is integrated over time with a leakage term following the LIF neuronal model. The model produces a spike when the integration or accumulation (despite the leakage) crosses a preset threshold. But, conventional LIF models can only handle unipolar input and emit spikes, on threshold crossing, without any labels. Thus, two copies of the model are normally needed to handle positive and negative parts of the input separately. In order to optimize the circuits and reduce area, we use a bipolar LIF (details in [Supplementary Note S1.](#)) having a shared integrator but with two thresholding comparators like the DM. Fig. 2b depicts the ON and OFF events generated by both DM and LIF when the input is a neural signal filtered to retain only the AP information. Both methods produce minimal spikes for background activity but emit a stereotypical set of spikes corresponding to the AP. However, the situation is very different when the ASC handles both LFP and AP in the shared channel approach described earlier. Using an added 10 Hz sinusoid with same amplitude as the AP to mimic the situation of added LFP (Fig. 2c), we see that the DM method produces sparse spikes to encode the slow LFP while still emitting the signature burst of spikes corresponding to the AP, as evidenced by the near perfect reconstruction of the AP. On the other hand, the LIF produces a very large number of spikes to encode the added LFP resulting in drastic loss of accuracy when reconstructing the AP. Given the importance of LFP, we also show that the slow signal can be reconstructed faithfully from the spikes produced by DM, with the AP reconstruction unaffected due to the timescale separation of the two types of spikes. Further, we perform a more detailed analysis of the AP reconstruction ability by varying the frequency (1-100 Hz) and amplitude (up to 10 times the AP) of the added sinusoid. Fig. 2e clearly shows far superior performance of the DM (normalized reconstruction error < 0.35) compared to LIF over the entire range of signals. Further, we see a trend of increasing error for both increasing amplitude and frequencies, as expected; but the errors remain quite small for DM in the entire range proving its superiority in tiny area ASC for future NNIs.

Finally, intracranial ECoG recordings often suffer from power line interference when the subject is close to electrical equipment or the mains line⁴⁵. This may result in signal to noise ratios (SNR) of as low as -20 dB, i.e. the interference can be ≈ 100 times larger than the signal. Removal of the interference is tricky since the signal is often non-stationary (due to subject movement, change in capacitive coupling etc.) and varies in amplitude and frequency. Many algorithms have been designed to handle this issue, and we demonstrate the results of applying one such algorithm⁴⁶ on the reconstructed signal for both DM and LIF spike encoders, using a publicly available 64-channel ECoG recording⁴⁷. Various noises (amplitude varying across a range of 40 dB and frequency varying from 45 - 65 Hz) were added to the original ECoG signal (details in Method) following the approaches adopted in an earlier work⁴⁶. As shown in Fig. 2f, the DM encoder consistently produces better noise removal properties (> 10 dB better SNR across all conditions) proving the superior signal separation capabilities of this ASC.

Bio-inspired Delta Modulation Schemes: Tradeoffs in Compression and Accuracy

We expect reduced datarate and concomitant compression from DM since it is expected to generate spikes only when a large signal (such as an AP) is observed at the input, while conventional Nyquist rate ADCs will generate samples all the time. However, the proper operation of DM requires the precise setting of a threshold to balance signal distortion and output data rate⁴⁹. If the threshold is too high, the reconstruction will be coarse, while if it is too low, noisy background signals will also generate a copious number of spikes. The timing of spike generation will also vary from the ideal one due to the presence of noise in neural recordings. In this situation, we look to the human brain for inspiration, which generates reliable behavior from noisy, and mismatched individual neurons. Among the different mechanisms used by biology, we focus on bursting neurons, found widely across the nervous system from neocortex and hippocampus in the brain to respiratory neurons and pancreatic β -cells. These neurons respond to input stimulus by firing a set of spikes in a burst (Fig. 3a) and these bursts may keep repeating, instead of the neuron firing a regular stream of spikes. The biophysical mechanisms of burst generation^{50,51} normally involve the interplay between a fast inward current (such as persistent sodium current, $I_{Na,P}$) that generate spiking activity and a slow outward current (such as persistent K^+ current) or a slow inactivation of the inward current (such as transient Ca^{2+} current). This slow process, gated either by voltage or Ca^{2+} concentration, eventually stops the burst. A simplified mathematical model (Izhikevich neuron, details in [Supplementary Note S2.](#)) is shown in Fig. 3a where the dynamics of the slow variable in the bottom define the burst period. Bursting neurons are believed to be very useful in neural computation⁵² by increasing reliability by overcoming synaptic failure, encode different features of sensory input, facilitate selective communication etc.

Here, we take inspiration from bursting neurons in the hippocampus, where there exist neurons with distinctly high threshold⁵³ leading to late bursts, as well as low-threshold early bursting ones⁵⁴. For our application, we postulate that having a high threshold is useful for reducing false spike generation. On the other hand, once this threshold is crossed signifying an important change in the input signal, switching to a lower threshold for a fixed duration will help to capture this event in more detail, thus providing a great balance between reconstruction error and datarate. Hence, we can mathematically define the

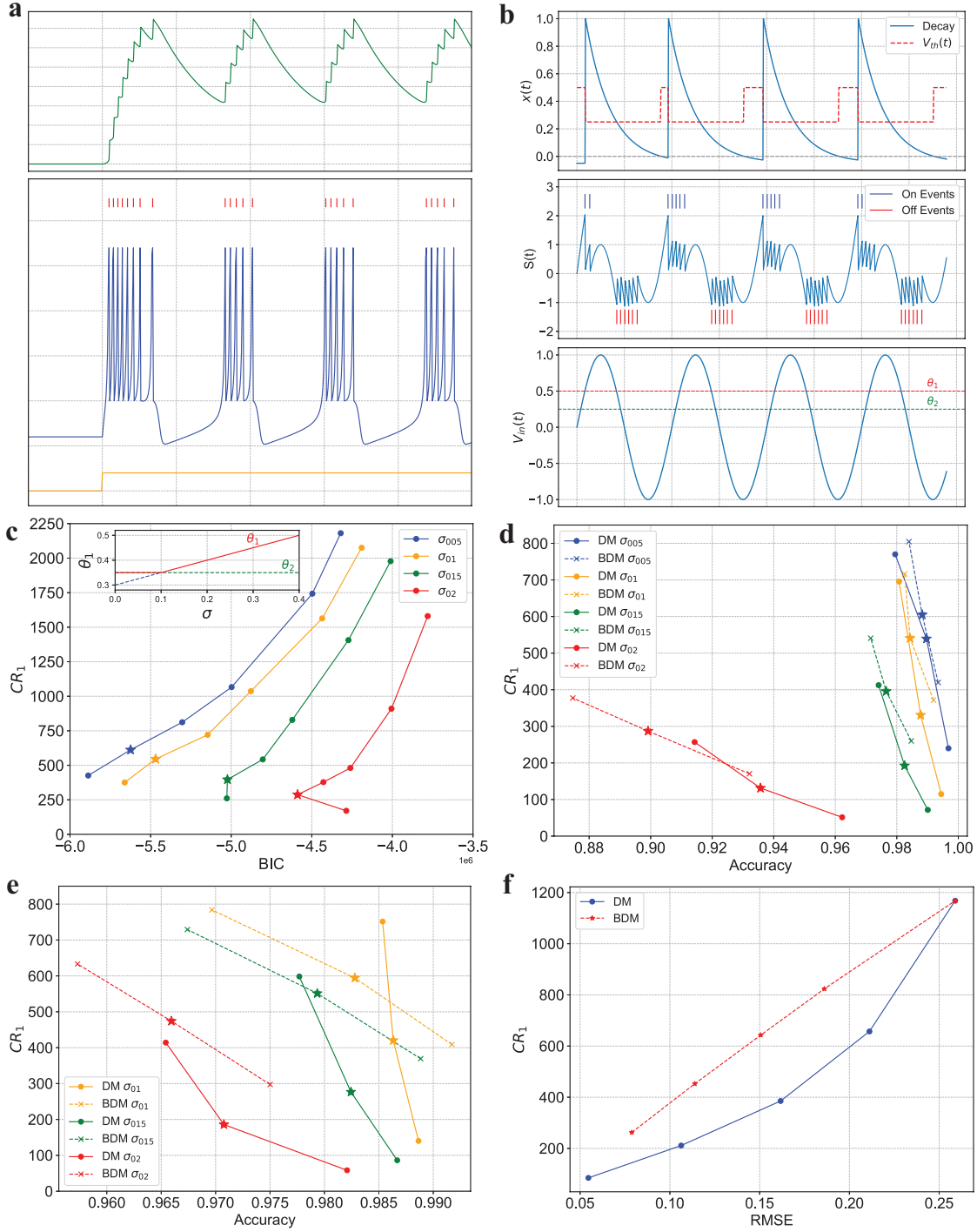


Fig. 3. **a** Bursting dynamics simulated by the Izhikevich neuron model is governed by the slow hyperpolarizing variable (top) which increases with every spike and eventually inhibits the neuron from generating further spikes. **b** BDM is an ASC inspired by the bursting dynamics. If the input exceeds the nominal high threshold θ_1 , the ASC switches to a smaller threshold θ_2 to emit spikes with accurate signal quantization. The duration of this burst is governed by a slow variable $x(t)$. **c** Bayesian Information Criterion (BIC) is used to analyze the rate-distortion tradeoff for BDM while varying threshold on the intracortical dataset⁴⁸ showing $k = 0.2$ (the $\star$ marker) being an optimal choice. Here, plot has different magnitudes of white noise as in the original dataset. **d** Comparison of BDM and DM's tradeoff between Compression Ratio (CR) and AP detection accuracy evaluated on the intracortical dataset⁴⁸ for $k = 0.1, 0.2$ and 0.3 . **e** Same as **d** but with added pink noise to mimic LFP signals. **f** Comparison of performance between DM and BDM with BDM using a fixed θ_1 (fixed at $k = 0.5$) with DM's threshold and BDM's θ_2 swept from $k = 0.1$ to 0.5 . BDM consistently achieve higher CR than DM ($\approx 2X$) at comparable threshold while achieving lower RMSE, thus dominating the pareto frontier.

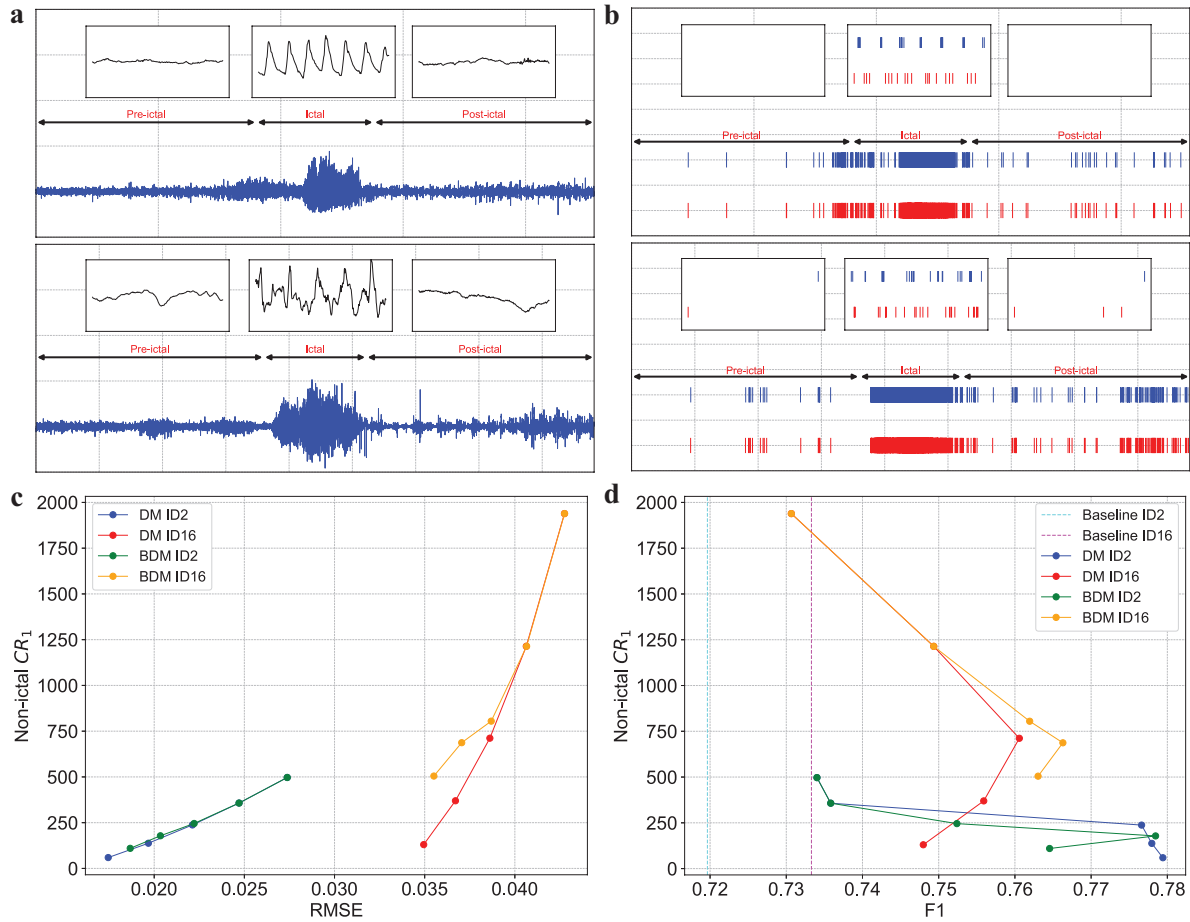


Fig. 4. **a** Example of iEEG signal for patient ID2 (top) and patient ID16 (bot) displaying three periods of pre-ictal, ictal and post-ictal data. **b** Example of spike train generated by DM encoding for patient ID2 (top) and patient ID16 (bot). **c** Comparison of BDM and DM's CR vs RMSE plot on the iEEG dataset when the threshold is varied. BDM consistently displays higher CR in the non-ictal periods which are the majority of the recording since seizures are very rare. **d** Comparison of BDM and DM's classification accuracies on the iEEG dataset showing comparable F1 scores, both slightly higher than the baseline, across the entire range of thresholds.

proposed Burst Delta Modulation (BDM) as follows:

$$S(t) = \begin{cases} 1, & V_{in}(t) - V_{ref}(t) \geq V_{th} \\ -1, & V_{in}(t) - V_{ref}(t) \leq -V_{th} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$\frac{dV_{ref}}{dt} = V_{th} \cdot S(t) \quad (2)$$

$$\tau \frac{dx}{dt} + x = -\varepsilon; \quad \text{If } |V_{in}(t) - V_{ref}(t)| > V_{th} \quad \text{and} \quad x(t) < 0 \implies x(t^+) = 1 \quad (3)$$

$$V_{th} = (1 - u(x))\theta_1 + u(x)\theta_2 \quad (4)$$

Here, Equations (1) and (2) define the conventional DM method where $V_{in}(t)$ is the input voltage, V_{th} is the threshold, $u(\cdot)$ denotes the Heaviside function and $S(t)$ is the output spike. The modulation of V_{th} by the slow variable is captured by the following equations, where Equation (3) defines the decay time of the slow variable and Equation (4) shows how the threshold switches between a large (θ_1) and a small (θ_2) one, where $u(x)$ denotes the Heaviside function. In between the ‘events of interest’, the slow variable x rests at $x = -\varepsilon$ (Equation (3)), and hence the threshold takes the large threshold value $V_{th} = \theta_1$. Whenever the difference in input and reference voltage (updated as Equation (2)) exceeds V_{th} , x is set to 1 and V_{th} switches to a smaller threshold value θ_2 for the duration required for x to decay exponentially below 0. The signal can be reconstructed from BDM events using the two thresholds and the burst state (BDM details in **Methods**) by either conveying an explicit flag denoting the state or implicitly inferring the state from the event stream. Fig. 3b shows an example when the input signal is a sinusoid. It can be seen that the burst of spikes starts only when the input crosses the large threshold θ_1 , while the generation of subsequent spikes is governed by θ_2 for a time window dictated by τ ($= 1$ ms for AP detection, for example) and ε (as defined in Equation (3)). Notably, the dynamics are quite similar to the ones depicted in Fig. 3a, with the exception that the biological neuron emits unipolar spikes.

There is a direct tradeoff between signal fidelity (defined by RMSE in **Methods**) and output data rate (defined by Compression Ratio (CR) with respect to conventional ADC in **Methods**)—a larger threshold results in more compression at the cost of poor signal reconstruction. We evaluate this tradeoff on the intracortical dataset⁴⁸ by sweeping the threshold (Supplementary Note S3.), where the range is normalized to the signal swing, i.e. $V_{th} = k \cdot \overline{V_{pp}}$ (see **Methods**). As can be seen from the inset, for BDM the larger threshold θ_1 was chosen based on standard deviation of background noise, while the fine threshold θ_2 was swept similar to DM based on signal swing (see **Methods** for details). This is also done for an augmented dataset (see **Methods**) where pink noise with a $\frac{1}{f}$ spectrum of different magnitudes are added to mimic the presence of LFP combined with AP. The magnitude of pink noise was also calibrated with real intracortical recordings¹⁵ using neuropixels probes such that the highest simulated noise matched those from the real recordings. To choose an optimal threshold from the tradeoff curves, we adapt the Bayesian Information Criterion (BIC) (see **Methods**) where the number of generated events is taken as the count of free parameters in the second term which trades off with the RMSE for AP reconstruction in the first term. The BIC is slightly modified from the original definition to a windowed BIC since we are only concerned with the RMSE in the AP regions. From the BIC plots in Fig. 3c for different levels of noise, it can be seen that $k = 0.2$ is optimal for white noise (similar plot in Supplementary Note S3. for pink noise cases).

While the CR in Fig. 3c varies from 1500 – 2000 depending on noise, the actual available compression may be lower for higher level tasks such as AP detection, as seen in recent rate-distortion-perception tradeoff theories⁴⁹. Accordingly, we train neuromorphic spiking neural networks (SNN) to evaluate the information output by both types of ASC for AP detection; this is done only for the optimal k obtained from BIC and its two closest thresholds ($k = 0.1$ and 0.3). A 2-layer SNN⁵⁵ was trained on a moving window of spikes emitted by the ASC (training details in **Methods**). A temporal voting of the raw SNN output was taken over a period of 1 ms to obtain the final classification result plotted in Fig. 3d and e for white and pink noise respectively. In all the cases, the accuracies and CR from the BDM are either the pareto frontier or extend the pareto frontier created by the DM (for example in the case of high white noise in Fig. 3d). CR values for BDM are consistently higher than DM, and range from 800 for low noise to 600 for high noise scenarios. Note that datarate of single channels are used to calculate CR_1 here, while practical packetizing methods⁵⁶ for multi-channel neural interfaces will incur some reduction in CR.

The earlier comparisons used carefully selected thresholds according to noise intensity; however, it becomes impractical in modern neural interfaces with thousands of channel where the noise power may also be non-stationary. In such cases, it is more practical to evaluate the performance of the ASC with fixed thresholds across varying types of signals. Hence, we evaluate the performance of BDM with a fixed coarse threshold of $\theta_1 = 0.5V_{pp}$ while varying the fine threshold θ_2 similar to that of DM. Fig. 3f plots the obtained tradeoff curve between RMSE and CR_1 when the lowest and highest noise power in the datasets were used for evaluation. The plot clearly demonstrating the superior performance of BDM ($\approx 2X$ higher CR_1 and lower RMSE for the same θ_2) across varying θ_2 , while at $\theta_2 = 0.5 = \theta_1$, DM and BDM's behavior become identical. Intuitively, this is because the higher value of θ_1 in BDM inherently reduces the number of spikes, while the switch to lower θ_2 preserves spike information.

To further evaluate the benefits of NNI, we select two patients from an intracranial EEG (iEEG) dataset^{57,58} with pharmacology resistant epilepsy recordings. Samples of the recordings are shown in Fig. 4a displaying 3 distinct regions of pre-ictal (before seizure), ictal (during seizure) and post-ictal (after seizure) zones. The goal for this case is to determine the ictal or seizure periods from the iEEG recording snippets. The recordings were first processed by both DM and BDM ASC techniques (Fig. 4b), demonstrating high compression ratios of $\approx 500 - 2000$ for the two patients, (Fig. 4c) in non-ictal periods which typically constitute $> 99\%$ of data since seizures are rare events⁵⁹⁻⁶¹. Similar to the earlier results, BDM extends the pareto frontier for DM to higher CR values of ≈ 500 and ≈ 2000 for patient ID2 and ID16 respectively. A SNN was also trained to classify the generated spikes (details in **Methods**) and the resultant F1 scores plotted in Fig. 4d demonstrate higher values than the floating-point baseline (with ADC based Digital Signal Processing). This was due to much higher false positives in the baseline, where the ASC methods excelled by generating very few spikes while encoding relevant features for the detection task.

Reconfigurable Burst Delta Modulator Integrated Circuit

We implemented a reconfigurable BDM integrated circuit in a 65 nm CMOS process, integrating it with a low-noise analog front-end (AFE) amplifier to make a complete front-end for NNI. The reconfigurable BDM (Fig. 5a) comprises a threshold-based delta-modulation encoder, a multiplexer for selecting between large and small thresholds, and a burst-control logic that dynamically selects between coarse and fine quantization windows similar to the slow variable dynamics in Equation 3. The reconfiguration from BDM to DM can be done by simply setting $\theta_1 = \theta_2$. The dynamics of the slow variable were implemented using a simple counter for ease of programming, and low-power/area implementation without needing large capacitors for long time constants. The AFE uses T-capacitor network in feedback to reduce area of the NNI pixel and implements a bandpass transfer function in the range of $\approx 0.1 - 5$ kHz. Details of the circuits are in [Supplementary Note S5](#).

The BDM hardware implementation was evaluated using neural recordings from the intracortical dataset, including 60-second segments of Easy2 and Difficult2, representing low- and high-noise conditions, respectively. The neural signals were input using a programmable waveform generator (see **Methods**), amplified and converted to spikes on the chip, and the generated ON/OFF event streams were recorded for analysis. Samples of the input waveforms and corresponding measured output spikes from BDM are shown in Fig. 5c (top and middle panels). To further evaluate robustness under low-frequency background modulation, a 10 Hz sinusoidal component with an amplitude $10\times$ the pulse baseline was superimposed onto the input signal to emulate low-frequency interference or LFP. The corresponding input waveform and measured BDM output are shown in the lower panel of Fig. 5b, demonstrating the robustness of BDM hardware event encoding under slow-varying drift conditions.

Based on the recorded event streams, we evaluated reconstruction accuracy and compression performance similar to the earlier simulation. Fig. 5d displays example of reconstructed AP waveforms using a zero order hold (ZOH) the measured spikes from the BDM integrated circuit. The different step size for the first event in a burst, corresponding to the larger threshold θ_1 is apparent here. The measured results exhibit CR of 10-25 while retaining spike shape information. The measured tradeoff between CR_1 and RMSE for BDM and DM is presented in Fig. 5e. As seen in software simulations (Fig. 3f), BDM displays a better tradeoff than DM, with a $CR_1 = 4130X$, while DM with similar RMSE can only obtain $CR_1 = 746X$.

Finally, we compare our work with other reported neural interfaces for BMI and in-vitro MEA usages in Table 1. Compared to other works, the proposed BDM achieves $120 - 300X$ ($\approx 40X$) higher compression than conventional ADCs (DM based ASC). The channel wise power consumption is also lower than other works with conventional ADC with similar input referred noise. Among the other ASC works, lower power consuming ones²⁸ had significantly higher noise as a tradeoff, or did not include the analog front end⁶², which consumes the largest share of power in our design.

Accelerated Analog Spiking Neural Networks for Rapid and Energy Efficient Event Classification

A major advantage of using ASC based NNI is the possibility of using low-power, event-driven SNNs in the implant to detect and transmit events of interest only, further reducing the transmission bandwidth manifold. As a proof of concept, we showed the possibility of using a full mixed-signal pipeline for AP detection by processing the outputs from the BDM IC using the BrainScaleS-2 (BSS-2) accelerator⁶⁸ (Fig. 6a, details in Methods). BSS-2 consists of 512 accelerated analog neurons modeling

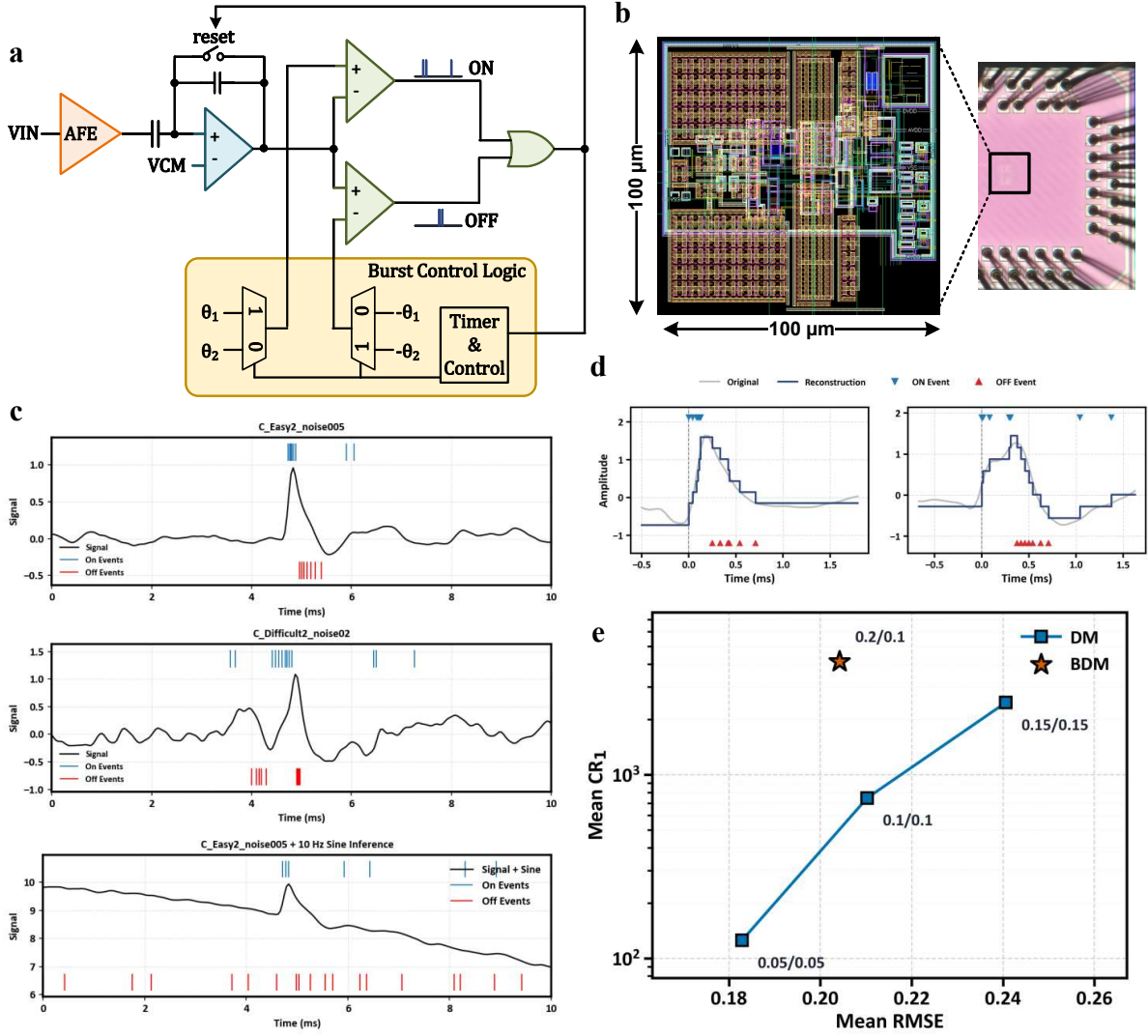


Fig. 5. **a** Schematic of the reconfigurable Burst Delta Modulator (BDM) circuit—setting $\theta_1 = \theta_2$ reduces the BDM to a conventional DM. The AFE is designed with a programmable gain of 500/1000 for AP detection. **b** Chip micrograph and layout of the BDM fabricated in 65 nm CMOS. **c** Measured spike output of the BDM integrated circuit for different noise levels and in the case of low-frequency interference, demonstrating the ability to generate sparse outputs. **d** Reconstructed waveforms from the BDM hardware at different noise levels. **e** Tradeoff between compression ratio and RMSE in BDM and DM measurements clearly show the superiority of BDM with $> 5X$ better CR_1 at similar RMSE. The value of θ_1/θ_2 is denoted for each point.

the Adaptive exponential and fire model and digital processors to program a wide variety of plasticity rules (unused in this work). Compared to other mixed-signal neuromorphic hardware^{69,70} that operate in real-time using sub-threshold currents and large capacitors, we chose an accelerated analog implementation to minimize area occupied by using smaller capacitors, and by sharing an AP detector SNN across multiple sensor channels resulting in an acceleration factor of > 20 for our SNN. Note that due to this accelerated operation of BSS-2, the SNN has to operate in a window mode and cannot operate in a streaming mode where the state needs to be retained individually for the AP detectors in each channel (details of window mode and streaming mode in [Supplementary Note S6](#).) Instead, in the windowed mode, buffers are needed to store only the ASC outputs for each channel of the MEA, which is then processed by the SNN. In our experiments, BSS-2 could process the SNN in a time step of $dt = 1 \mu\text{s}$, resulting in an inference time of $T_{inf} = 24 \mu\text{s}$ per AP detection by processing events in a window of $T_{history} = 1 \text{ ms}$ using step of $S_{win} = 1$ timestep within the window. The final output is obtained by temporal voting across all the timesteps, as explained in the earlier sections. However, the net acceleration of the SNN depends on the stride S when moving to the next window as well, and is given by $\lfloor \frac{S \cdot T_{hist}}{T_{step} \cdot T_{inf}} \rfloor$

Tab. 1. Comparison of the proposed BDM with other reported analog-digital converter hardware in implantable BMI systems.

Publication	Tech. [nm]	Pixel size [μm^2]	ADC/ASC	Data rate/ch. [kbps]	Power/ch. [μW]	IRN [μV_{rms}]	Application
Nat. Comm.'24 ²⁸	180	48×48	DM	–	0.8	19.05 (0.5–3 kHz)	MEA
TBCAS'24 ⁶²	65	326 (w/o AFE)	DM	39.1	1	–	BMI
JSSC'15 ⁶³	65	36300	SAR	200	3.59	7.5 (10–8 kHz)	BMI
ISSCC'24 ⁶⁴	65	9000	SAR	200	3.44	8.8 (0.3–10 kHz)	BMI
Nat. Comm.'17 ⁶⁵	180	25×25	SAR	120	–	26 (100–10 kHz)	MEA
JSSC'18 ⁶⁶	130	15×15	SAR	300	19.8	7.5 (100–3 kHz)	MEA
JSSC'17 ⁶⁷	180	13.5×13.5	SAR	200	42	2.5 (300–10 kHz)	MEA
This work	65	100×100	DM+BDM	1.01	2.85	7.56 (200–7.2 kHz)	BMI

The trained SNN described earlier was mapped to hardware. The inevitable mismatch induced errors were tackled by a special chip-in-the-loop training where forward passes on the BSS-2 hardware were followed by backward gradient calculations in software⁷¹ (Methods). The output events from the BDM IC (for a low-noise and high-noise case) were stored in a file and were separately input to BSS-2 using an API over the internet. Fig. 6b shows an example of the raw input, BDM outputs and final outputs from the two positive detection (P) and negative detection (N) neurons on BSS-2. The difference of P and N accumulated over a window determines the final classification result. Comparing P-N with a varying threshold T produces a precision recall curve (Fig. 6c) showing excellent performance. In use cases, the value of (P-N) may be transmitted (using ≈ 2 bits per spike) along with detection results for further classification off the implant. The classification performance of the fully analog pipeline is compared with the baseline in Fig. 6d. For both cases, the result obtained from BSS-2 is nearly identical with the SNN software that acts on the IC output of BDM, while reducing the data rate compared to BDM by 37/13X for the cases of high and low noise, respectively. Two cases for BSS-2 measurements are shown: one with stride of $S = 1$ that has higher accuracy but no acceleration, and another with stride of $S = 12$ (implying 74% overlap across windows) that results in an acceleration of ≈ 20.8 . This implies that the SNN can be shared among 20 channels resulting in a significant saving of silicon area.

Discussion

Traditional multi-channel neural interfaces employing a conventional clocked sampling scheme have limited scalability due to fundamental limitations of bandwidth and power dissipation resulting from increasing datarates. The problem is exacerbated for implantable microelectrode array based BMIs¹⁵, where long-term usage requires wireless transmission links for patient safety and comfort. Similar constraints affect long-term ambulatory recording of iEEG in patients suffering from neurological conditions such as drug resistant epilepsy. As a result, current devices either provide a small snapshot of brain recordings when stimulation is triggered⁷², or longer term recordings from only a few channels^{73,74}. To break this tradeoff, in this work, we proposed and experimentally demonstrated an analog neuromorphic neural interface that replaces a conventional analog to digital converter with a bio-inspired analog to spike converter (ASC) that takes advantage of the natural sparsity of desired biological events of interest to provide manifold compression at source. We demonstrate the operation of our method on two extremely different detection tasks—detecting short, millisecond level action potential events as well as slower second level dynamics of epileptic seizures.

We experimentally demonstrated the Delta modulator (DM) class of ASC to be robust against interference signals (> 10 dB improvement over LIF models), common in bio-potential recording. It should be noted that the LIF neuron models are also closely related to Delta Sigma ADCs⁷⁵, popular in high resolution sensor interfaces. While conventional circuit design aims to increase the resolution of a single such ADC, note that our work goes in the opposite direction and effectively reduces resolution of a single ADC while achieving high net information transfer due to the use of thousands of such circuits. Also, low-channel count interfaces such as ECG or EMG have used more sophisticated LIF encoders with pre-filtering, adaptation etc. and suffer far less from interference issues. However, their large footprint makes them unsuitable with high density neural interfaces.

Further, we presented a new class of DM inspired by hippocampal large threshold bursting neurons that provide a much better tradeoff between data compression and signal recovery ($\approx 5X$ better compression than DM at same fidelity). The proposed BDM presents an excellent compromise between circuit complexity, compression and reconstruction/detection of events of interest. This was experimentally validated by using an analog spiking neural network that could process the generated spikes to detect neural action potentials within 2% of software accuracies, while reducing datarate after detection by 13 – 37X, paving the way for high density neural recording systems in future.

A shortcoming of our current work is that system level serial communication overheads⁵⁶ such as packet headers, clock recovery from data etc. are not considered. While the compression rates reported here may be an upper limit for the system

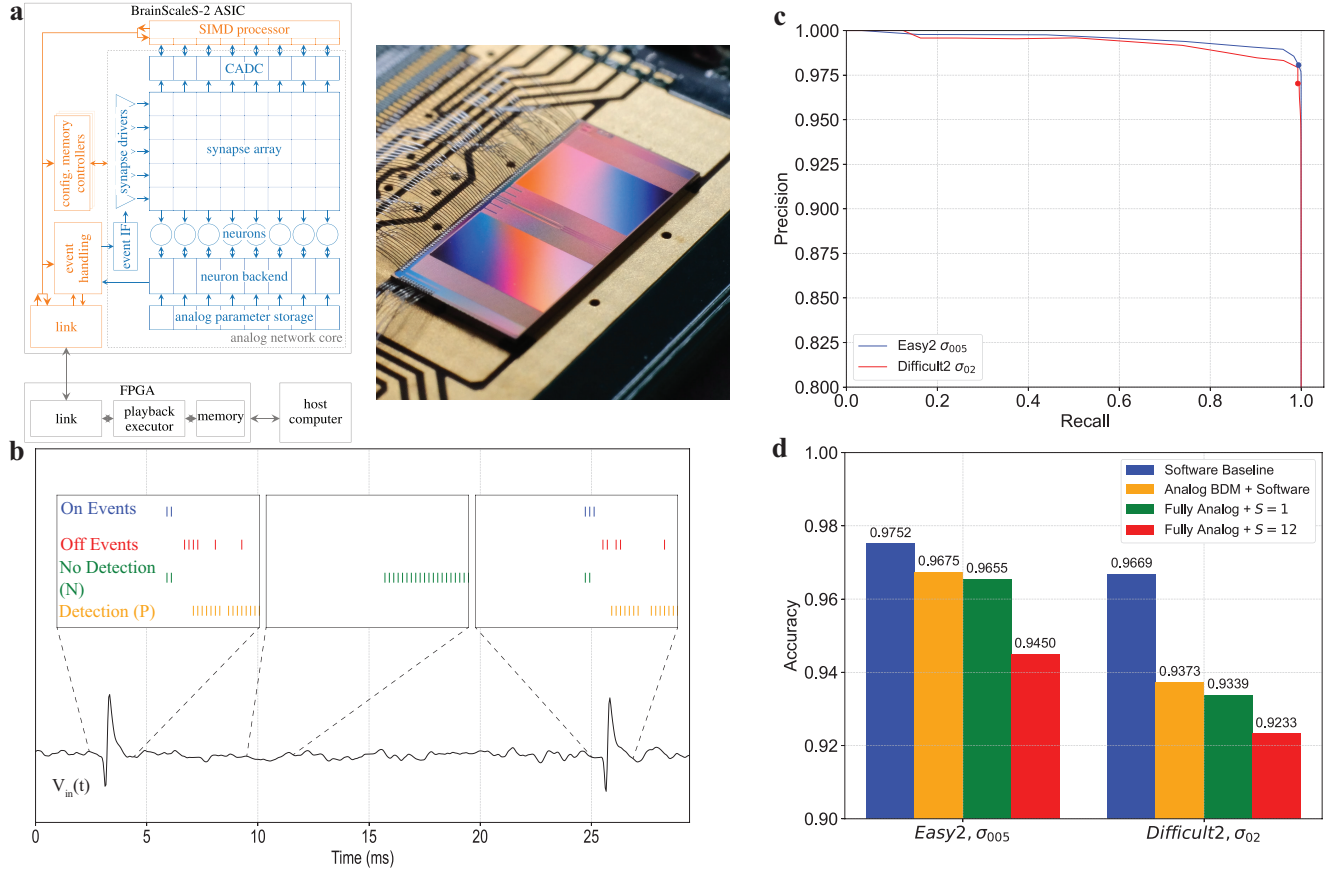


Fig. 6. a BrainScaleS-2 die photo and block diagram⁶⁸. **b** The input neural signal $V_{in}(t)$ is converted into spike train $S_{in}(t)$, which is then passed into the SNN model hosted on the BrainScaleS accelerator for inference. For AP detection, the spikes of positive detection (P) and negative detection (N) neurons are summed within the detection window and the neuron with the highest firing rate determines whether AP is detected. **c** Precision-Recall plot of AP detection on BrainScaleS obtained by sweeping a threshold T compared with P - N. The points with $T = 0$ are marked. **d** Comparison of the efficacy of AP detection for $T = 0$ across three different setup, full software simulated baseline, analog BDM IC front-end with software simulated inference and fully analog pipeline (analog front-end with BSS-2 accelerator for inference) on two files with two vastly different noise levels. BSS-2 achieved nearly identical performance as software.

level performance, our work goes beyond just the ASC and defines a new paradigm where conventional data sampling is replaced by computational, task aware sensing. Placing such analog to information conversion in the early sensing stages enables to break the strict tradeoff between increasing sensor density and managing the concomitantly burgeoning datarates. We envision future neural interfaces to increasingly adopt such task aware sensing that optimize downstream task performance rather than focusing entirely on increasing the fidelity of the captured signal, helping create smart, closed-loop systems for motor prosthesis, in-vitro neuroscience experiments, epileptic seizure treatment and much more.

Methods

This section outlines the theories and experimental details used in this work.

ASC Software Simulation Equations

Below is a list of standardised variables used in this project and the corresponding definition.

- $V_{in}(t)$: Input Voltage Signal
- $V_{th_{on/off}}$: The On/Off threshold Voltage
- $V_{ref}(t)$: Reference Voltage

- $S(t)$: Spike Train
- $\overline{V_{pp}}$: Median Peak-to-Peak Voltage
- V_{recon} : Reconstructed Signal
- τ : Time Constant
- k : Multiplier/Constant swept (if applicable)

Compression and Reconstruction Process

To compare the performance of the ASC encoders, we used the root mean square error (RMSE) and compression ratio (CR) to quantify the reconstruction error and data compression, respectively. In this work, reconstruction error refers to the reconstruction of the input signal $V_{in}(t)$ from the generated spike train $S(t)$. In our evaluation, the RMSE is defined according to Equation (5), where N_s is the total number of samples in $V_{in}(t)$, $W(t)$ is a window function with magnitude equal to 1 for T_{spk} ($=1$ ms) around the AP ground truth times and 0 elsewhere, while $N_W = \frac{T_{spk}}{T_s} \cdot f_{neu}$ denotes the total number of samples in these windows, where f_{neu} is the total number of AP in the recordings.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_s} W(iT_s)(V_{recon}(iT_s) - V_{in}(iT_s))^2}{N_W}} \quad (5)$$

CR is defined as the ratio of data rate in conventional ADC and the proposed ASC. Considering address event representation (AER) being used to encode the address of a channel in a multi-channel probe, we rename CR as $CR_{N_{channel}}$ as shown in Equation (6), where N_{spikes} is the number of non-zero events in $S(t)$, N_s is the number of samples, b_{adc} is the ADC bit resolution (in our work $b_{adc} = 12^{76}$) and $N_{channel}$ is the number of recording channels.

$$CR_{N_{channel}} = \frac{N_s \cdot b_{adc}}{N_{spikes} \cdot (\lceil \log_2(N_{channel}) \rceil + 1)} \quad (6)$$

Note that this method is assuming the transmission of spikes as soon as they occur. It has been shown that this is not necessary for NNIs and creating sparse event count frames at ≈ 20 kHz is enough for high performance⁷⁶.

Bayesian Information Criterion (BIC)

In order to determine the ideal k (or equivalently V_{th}) for the different ASC encoders, we used the Bayesian Information Criterion (BIC)^{77,78}, a commonly used information theoretic criteria for selecting optimal model based on trade-off between model error rate and model complexity. BIC is defined according to Equation (8)⁷⁷, where T_s is a sampling time resulting in N_s total number of samples, κ is the model complexity/number of non-zero parameters and the remaining symbols are as defined earlier. In the context of ASC, κ denotes the number of generated events.

$$BIC = N_s \log\left(\frac{\sum_{i=1}^{N_s} (V_{recon}(iT_s) - V_{in}(iT_s))^2}{N_s}\right) + \kappa \log(N_s) \quad (7)$$

It can be seen that the two terms capture the involved tradeoff—the first term reduces if reconstruction error is low, while the second term reduces if the datarate due to number of generated events is low. However, for our case, the equation needs to be modified (or weighted appropriately) since the reconstruction error *only* in events of interest (e.g. AP in intracortical dataset) are important—rmse in other areas are redundant. Hence, we modify the basic definition of BIC as follows:

$$BIC = N_s \log\left(\frac{\sum_{i=1}^{N_s} W(iT_s)(V_{recon}(iT_s) - V_{in}(iT_s))^2}{N_W}\right) + \kappa \log(N_s) \quad (8)$$

The definition of $W(t)$ and N_W is the same as in the previous subsection.

Delta Modulation

Delta Modulation is a classic technique used for analog-to-digital conversion, where the difference of the input voltage $V_{in}(t)$ and a reference voltage $V_{ref}(t)$ is compared with a threshold to generate spikes or events. Generally, we can define $V_{th_{on}}$ as the ON event threshold and $V_{th_{off}}$ as the OFF event threshold, but for simplicity we use $V_{th} = V_{th_{on}} = V_{th_{off}}$. We define a positive change exceeding V_{th} to be an ON event/spike, and a negative change to be an OFF event/spike, as seen in Equation (1). When an event is generated, the reference voltage V_{ref} is updated by Equation (2).

In practice, the threshold V_{th} can be defined with respect to the signal swing. For example, if there is variability in the peak to peak of the signal to be detected, we can define $\overline{V_{pp}}$ as the median peak-to-peak voltage of the input signal, and the threshold can be obtained through Equation (9) where the scaling parameter k determines the value of V_{th} . As k increases, the number of events generated will decrease proportionally, increasing the compression ratio and RMSE of the reconstructed signal.

$$V_{th} = k \cdot \overline{V_{pp}} \quad (9)$$

In our work, we swept a range of values for the multiplier k with the goal of finding the tradeoff between RMSE and CR. In reconstruction, zero order hold was used to create a staircase approximation, unless otherwise mentioned.

Burst Delta Modulation

Burst delta modulation (BDM) is a dual-threshold extension of conventional delta modulation. It is designed to simultaneously suppress background noise and preserve transient high-frequency neural activity. The equations for this method are described in the main text with two thresholds denoted as θ_1 and θ_2 .

$$\begin{aligned} \theta_2 &= k_2 \cdot \overline{V_{pp}}, \\ \theta_1 &= \max(A + k_1 \sigma, \theta_2). \end{aligned} \quad (10)$$

The two thresholds are estimated from both noise and spike statistics. The smaller threshold θ_2 is proportional to the median peak-to-peak amplitude $\overline{V_{pp}}$, which ensures that burst-mode sampling remains matched as DM. The larger threshold θ_1 is determined by a baseline shift A normalized to the amplitude and the background noise level σ . The constraint $\theta_1 \geq \theta_2$ guarantees that the nominal mode is no more sensitive than the burst mode. The duration of small threshold governed by Equation (3) is implemented as a fixed time count in software code and as a programmable counter in hardware. Similar to DM, k was swept to determine the tradeoff between RMSE and CR.

The reconstruction of neural signal from BDM events requires knowledge of the active threshold state associated with each event. This information can either be explicitly conveyed by an additional one-bit flag from the modulator or implicitly inferred at the receiver from event timing, since the burst-state transition is deterministic. The first event within a burst window is reconstructed using the coarse threshold, while subsequent events within the same window are reconstructed using the fine threshold. The second method of implicitly determining the threshold was used in this work. In reconstruction, zero order hold was used to create a staircase approximation, unless otherwise mentioned.

ASC Hardware Simulation

All the ASC circuits were simulated in TSMC 65 nm CMOS process.

Delta Modulation

The DM ASC circuit was evaluated by transistor-level SPICE simulations. The DM encoder converts the amplified neural waveform into bipolar ON/OFF events through threshold-triggered comparators. After each event, the reference node is updated to continue tracking the input waveform. We conducted a neural signal dataset-driven simulation. The neural signal is converted into voltage inputs applied to the delta-modulator circuit. The generated ON/OFF pulses were recorded as the spike output for ASC performance analysis. The DM circuit was biased with a total current of 1.15 μ A. Under dataset-driven stimulation, it consumed 1.38 μ W from a 1.2 V supply and occupied 1728 μ m² per channel. Transistor sizes and layout of the DM circuit are provided in [Supplementary Note S5](#).

Burst Delta Modulation

Compared with conventional DM, the burst delta-modulation (BDM) encoder uses a dual-threshold scheme to balance noise suppression and signal recovery. A nominal coarse threshold window is used to suppress noise-induced transitions. Once a coarse event is detected, the circuit switches to a fine-threshold window for a short duration, allowing burst events to provide finer tracking of neural transients. Apart from the delta-modulation core, the BDM circuit incorporates an additional burst-control path to switch between two effective thresholds. For dataset-driven simulation, neural signal traces were applied to the BDM circuit as voltage inputs, and the generated event polarities and intensities were recorded as the ASC spike output for performance analysis. The BDM circuit was biased with a total current of 1.17 μ A. Under dataset-driven stimulation, it consumed 1.4 μ W from a 1.2 V supply and occupied 1920 μ m² per channel. Transistor sizes and layout of the BDM circuit are provided in [Supplementary Note S5](#).

ECoG Dataset and Removal of Power Line Interference

We used real ECoG signal from the ECoG⁴⁷ dataset, which contains recordings of two Japanese macaques monkey fitted with a 64-channel ECoG channels in the epidural space of the left hemisphere covering the area from the prefrontal cortex to the primary somatosensory cortex. The recordings are of the monkeys moving hands to track food rewards sampled at 1 kHz. To simulate the effect of power line interference, the ECoG signal is tempered with synthetic noise with three harmonics at 61 Hz, 122 Hz and 183 Hz. The tempered ECoG signal is then passed through a DM/LIF encoder to generate an event stream, which is then used to reconstruct the tempered signal. With the reconstructed signal, the power line interference can be removed using the algorithm proposed by Keshtkaran et al⁴⁶. With this algorithm, we first apply a 4th-order infinite impulse response (IIR) band-pass filter between the band 40 Hz and 70 Hz to the reconstructed signal to enhance the fundamental harmonic of the interference signal. To reduce estimation bias, a 1st-order differentiator is used to mitigate the effect of the power law spectrum. We then use a lattice adaptive notch filter (ANF)-based frequency estimator to estimate the interference fundamental frequency, $\hat{\omega}_f$. Based on $\hat{\omega}_f$, the harmonic signals $h_k(n)$ can be obtained with discrete time oscillators. With $h_k(n)$, the amplitude and phase of the harmonic signal can be estimated with recursive least squares (RLS) algorithm. This will generate the final estimated interference signal, which is subtracted from the reconstructed signal to generate the clean signal.

We set up two experiments to illustrate the difference in performance between DM and LIF ASC encoder. The first experiment varies the strength of the synthetic noise at a fixed sampling rate of 1,000 Hz, where we compared to SNR of the tempered signal (SNR_{in}) against the SNR of the output signal (SNR_{out}). SNR_{in} was varied from -20 dB to 20 dB. The second experiment varies the fundamental harmonic frequency between 45 to 65 Hz and evaluated the SNR_{out} for performance. For both set of experiments, we ran each setting 50 times.

Intracranial EEG Dataset for Seizure Detection

To evaluate the performance of the ADM and BDM ASC methods in a multi-channel setup, we utilized an anonymized intracranial EEG (iEEG) dataset. These data were obtained from a public epilepsy dataset^{57,58}, assembled by the Sleep-Wake-Center (SWEC) of the University Department of Neurology at the Inselspital Bern (available at <http://ieeg-swez.ethz.ch/>). The dataset comprises recordings from multiple patients with pharmacoresistant epilepsy, collected during pre-surgical evaluations using multi-channel strip, grid, and depth electrodes placed according to the hypothesized location of epileptogenic tissue. The iEEG signals were sampled at 512 Hz and pre-filtered within the 0.5 to 150 Hz frequency band using both forward and backward filtering. Channels containing artifacts were removed during pre-processing. Seizure onsets and offsets were provided as annotations by a board-certified, experienced epileptologist. For each patient, multiple seizure events were recorded, each consisting of a three-minute pre-seizure segment, a seizure segment (ranging from 10 to 1002 s), and a three-minute post-seizure segment.

Due to computational constraints, we selected a subset of the dataset containing two patients (ID2 and ID16). Patient ID2 has recordings from 42 electrodes across four seizure events, while Patient ID16 has recordings from 59 electrodes across six seizure events. Data from the ictal periods were labeled as 1, and non-ictal periods (including the aforementioned pre- and post-seizure segments) were labeled as 0. The objective of this task is to identify the specific sections of the recordings in which the patient is experiencing a seizure event.

Intracortical dataset for AP detection

To comprehensively evaluate various ASC encoders for AP detection with SNN, we selected a well known synthetic intracortical dataset⁴⁸ and extended it. This is a synthetic dataset constructed from a database of 594 spike shapes with their amplitude normalized to have a peak of 1. The dataset comprises of files labeled based on their difficulties and noise levels (standard deviation of background noise relative to the spike amplitude), with each file having a duration of 60 seconds. Within each files, there are varying number of spikes, ranging from 3,364 to 3,526.

The background noise in the intracortical dataset is white Gaussian noise. The dataset provides files with varying difficulty and are named *Easy1*, *Easy2*, *Difficult1* and *Difficult2*. As for noise levels, 4 different ones are provided— σ_{005} , σ_{01} , σ_{015} and σ_{02} , thus giving us 16 combination of files to train and test on. Since files with noise level σ_{02} contain SNR levels found in real world dataset such as Neuropixel¹⁵, this would ensure our analysis remain grounded with realistic setting.

Since the original intracortical dataset did not contain LFP, we extended it by creating a set of files contaminated with $1/f$ noise (pink noise) to mimic LFP added with AP. We used noise level σ_{005} of difficult *Easy1*, *Easy2*, *Difficult1* and *Difficult2* as the template. We first generate white noise than inverts it to obtain $1/f$ noise (pink noise). The strength of the pink noise is varied such that the SNR of the final signal tempered with pink noise matches the SNR of the corresponding white noise file of identical difficulty and noise level, i.e. the pink noise *Difficult1* at noise level σ_{02} have the same SNR level as the white noise *Difficult1* at noise level σ_{02} .

SNN training for AP detection

Pre-processing: Each of the recording is first passed through a first order band-pass filter between the frequency range of 1 Hz and 5,000 Hz. After filtering, the signal is passed into the chosen ASC to generate the spike train $S(t)$. The spike train is split into two, where the first 80% is used for training while the remaining is used for testing. Due to the significant class imbalance as APs are sparse in nature, during training, the spike trains are split into windows for both AP and non-AP, and an equal number are selected. The dataset have provided the ground truth timestamp for when spike occurs, which is delayed by 1.5 ms.

Architecture: In this work, we designed our model based on the work of Hwang et al⁵⁵, where a two fully-connected layers SNN with LIF spiking neurons of dimensions 16 and 2 is used. The SNN operates in two modes, window and streaming. During training, the SNN is trained using window mode, where the spiking neurons reset at the start of each prediction. In window mode, each window L has 47 time steps. In our simulations, for inference, the SNN operates in streaming mode, where the spiking neurons do not reset. For both modes, the SNN expects an input segment of 24 consecutive time steps. A full summary of the setup of the SNN model can be found in [Supplementary Note S6](#). To enable training through backpropagation, the LIF neurons use the arctan surrogate gradient function⁷⁹ and reset with the reset-by-subtraction mechanism, in order to retain prior information.

Training method: Our model is trained with the Adam optimizer with a learning rate of 0.0005 for 10 epochs, and the loss function used is the spike count cross entropy loss, where we optimize for the spiking neuron that corresponds to the correct class to fire frequently. For training, we selected the first 80% of ground truth spikes of each file for training/validation and the last 20% for testing. We train on all 16 files for the white noise intracortical dataset and all 12 files for the pink noise intracortical dataset, so we only train one set of weights for each k of the ASC encoder. The accuracy metric for AP detection is defined by Equation (11)^{55,80}. During AP detection, if the SNN was to output spikes within 0.5ms of the ground truth time, we classify this as TP.

Inference operation: During inference, the SNN's classification process is based on temporal voting, where the output neuron with the most spikes generated determines if AP is detected. For SNN in window mode, the temporal voting is done for the duration of the window, while in streaming mode, the temporal voting is calculated over a 1ms period.

$$Acc = \frac{TP}{TP + FP + FN} \quad (11)$$

SNN Training for Seizure detection

Architecture: We designed an SNN model comprising three fully-connected layers with LIF spiking neurons of dimensions 177, 118 and 1. The layers' dimensions and LIF neuron's decay rate and firing thresholds are chosen via parameter sweep with the goal of maximizing detection accuracy. Similar to the SNN model used for AP detection, this SNN model operates in two modes, window and streaming. During training, the SNN operates in window mode, expecting N_{ch} inputs (which is determined based on the patient ID) for 32 time steps, while in the testing phase, the SNN operates in streaming mode, expecting the same input dimension as in window mode. The LIF neurons operate in the same manner as the LIF neurons used in , except for the surrogate gradient, which uses the Double Gaussian surrogate gradient function⁸¹ instead.

Training method: The AdamW optimizer is used to train our model, with a learning rate of 0.001 for 20 epochs. The binary cross entropy loss is applied to the output of the SNN and is encouraged to fire only in the ictal periods of the recordings. For each patient, we split the recordings according to the leave-one-out principle, where $n - 1$ recordings are used for training and validation and 1 recording is left for testing. The final reported F1 score is based on the average result of the leave-one-out training, and the F1 score is defined according to Equation (12), where TP , FP and FN refer to true positive, false positive and false negative, respectively. In order to showcase the efficacy and benefits the ASC encoders have on the SNN, we performed a baseline result using the same SNN model parameters, and changed the input from $S(t)$ to $V_{in}(t)$, i.e. using the values taken from the recording are directly input as currents to the first layer of neurons in the SNN.

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (12)$$

BDM Integrated Circuit Measurement

We performed hardware testing of BDM ASC on an integrated circuit as shown in Fig. 5. The chip under test was mounted on a custom printed circuit board, which is connected to a Keysight 33500B waveform generator via an SMA port and to a ZYNQ7015 field-programmable gate array (FPGA) board via an FMC socket (as in [Supplementary Note S7](#).). The synthesized test neural inputs are transmitted to the waveform generator and then applied to the BDM circuit via the waveform generator. We observed transient waveforms from the AFE and the event outputs from the BDM through a Tektronix TBS2000B oscilloscope and measured the power dissipation of the IC with a Keithley 6485 picoammeter. The event outputs from the chip are captured

and timestamped by the FPGA board operating at a clock frequency of 50 MHz, and the timestamp and polarity of the events are sent to the PC for evaluation. To align the timestamps with the start of each neural input, a short triangular-wave trigger segment was added before every input waveform. The trigger consisted of 10 cycles with a period of 1 ms and was used as a timing marker. After acquisition, the event sequence generated by this trigger segment was identified from the raw FPGA timestamps, and the end of the trigger segment was defined as the onset of the neural input. Subsequently, we calculate the compression ratio based on the rate of events in the experiments and derive the RMSE by comparing the reconstructed waveform with the original waveform.

SNN Implementation on BrainScaleS-2

In order to showcase the viability of a fully analog pipeline for AP detection, we setup an experiment to compare the detection performance between software simulated AP detection and hardware AP detection. In this experiment, we used the same model architecture and training method described in the **SNN training for AP detection** subsection and [Supplementary Note S6](#). We setup three different experiments to showcase the viability of a fully analog pipeline. First, we establish a software baseline where the BDM frontend and SNN inference are both simulated. Next, a hybrid setup where analog BDM IC is used to generate the spike train with simulated SNN inference. Finally, a fully analog setup where analog BDM IC is used for spike train generation and SNN inference on the BrainScaleS-2 accelerator.

In our previous experiments, the SNN model operated in streaming mode during inference; however, on hardware with multichannel input, there are resource constraints making this operating mode infeasible. Window mode inference, on the other hand, allows resources to be shared across all channels, as the membrane potentials of the LIF spiking neurons in the SNN do not need to maintain membrane potential in perpetuity, which streaming mode requires. Therefore, on the BrainScaleS-2 accelerator (and in our software simulation for fair comparison), the SNN model uses the window mode instead. To maximize the acceleration that can be obtained with the BrainScaleS-2 accelerator, we applied striding to the inference process. During testing, the SNN's inference window is shifted forward by S time steps, and S can be varied to obtain a tradeoff between accuracy and resource constraint. As mentioned earlier, temporal voting is applied during inference to determine if AP is detected.

In the hybrid setup and full analog setup, the spike train is generated using our BDM IC, using difficulties *Easy2* and *Difficult2* at noise level σ_{005} and σ_{02} , respectively, which covers both difficulty and noise level at both extremes, while keeping the test time manageable.

To enable the transfer of our SNN model from software simulation to real hardware accelerator, we mapped the software parameters in Equation (13) describing the discrete LIF to hardware parameters in Equation (14) describing the adaptive exponential integrate and fire model. To bridge the gap between the SNN model in software and physical deployment, we transition from `snnTorch` to the BrainScaleS-2 neuromorphic hardware using the `hxtorch` API⁷¹, which provides a time-gridded compute graph abstraction over the chip's time-continuous dynamics. The BrainScaleS-2 architecture relies on mixed-signal physical emulation, which accelerates neural dynamics by a factor of 1000 compared to biological real time. Under this (configurable) acceleration, model time constants typically spanning milliseconds are physically emulated on the hardware at the scale of microseconds. This drastic compression of the execution timescale enables rapid physical inference. To mitigate inherent physical distortions and device mismatches on the analog neuromorphic substrate, `hxtorch` provides a Hardware-in-the-Loop training framework. In this setup, the forward pass is executed directly on the physical accelerator to capture real hardware dynamics, while the backward pass is executed numerically in software to calculate gradients. By injecting authentic hardware responses directly into the optimization loop, the network parameters adapt to the specific nuances of the silicon substrate, enabling high-accuracy inference while minimizing hardware resource constraints.

$$\begin{aligned}
\beta &= e^{-\delta t / \tau} \\
U[t+1] &= \beta U[t] + W S_{in}[t] + \theta[t] \\
S_{out}[t] &= \begin{cases} 1, & U[t] > U_{th} \\ 0, & \text{otherwise} \end{cases} \\
\theta[t] &= S_{out}[t-1] U_{th}
\end{aligned} \tag{13}$$

$$\begin{aligned}
C_m \dot{V} &= -g_1(V - E_1) + g_1 \Delta_T \exp\left(\frac{V - V_T}{\Delta_T}\right) - w + I \\
\tau_w \dot{w} &= a(V - E_L) - w
\end{aligned} \tag{14}$$

Data availability

The data supporting plots within this paper and other findings of this study are available with reasonable requests made to the corresponding author. The public datasets used in the article are sourced from the following :

- The ECoG dataset can be found at <https://neurotycho.org/epidural-ecog-food-tracking-task>.
- The iEEG dataset can be found at <http://ieeg-swez.ethz.ch/>.
- The intracortical dataset can be found at DOI: [10.25392/leicester.data.11897595](https://doi.org/10.25392/leicester.data.11897595)
- The extended intracortical data-files with added pink noise are hosted at DOI: [10.5281/zenodo.21215958](https://doi.org/10.5281/zenodo.21215958)

Code availability

The code used to train the model and perform the simulation on AP detection and seizure detection is publicly available in an online repository [BNEW2025-26 ASC Project](#) in Github.

Acknowledgments

The project was conceived and partially performed during the Bangalore Neuromorphic Engineering workshops (BNEW) of 2025 and 2026.

The work of A.B, P.S.V.S, Y. K, Z. F, A. T and Z.B was supported by RGC of the HKSAR, China (Project No. CityU 11200922 and CityU 11208125).

S.G. acknowledge Swiss National Science Foundation for project TMPFP2_217160 for partially supporting this work.

This manuscript has been authored in part by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-publicaccess-plan>). This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under contract number DE-AC05-00OR22725.

Q.H. acknowledges the financial support of the VolkswagenStiftung (Volkswagen Foundation) through the SNACE project. The BrainScaleS-2 platform used for SNN hardware experiments was co-funded by the EC Horizon Europe Framework Programme under grant agreement 101147319 (EBRAINS 2.0).

Author contributions statement

A.B. conceived the idea and coordinated the project with S.G. P.S.V.S performed DM simulations and SNN training in software and hardware. Y.K. performed BDM simulations, co-designed the chip with Z.F. and performed measurements with Z.F. B.Z. performed simulations involving LIF. C.D.L. helped with pink noise simulations. E.A helped with the SNN hardware experiments on Brainscales and helped write the corresponding section with E.M and J.S. P.S.V.S, Y.K, S.G. and A.B wrote the manuscript with inputs from all authors.

Apart from the first two and last five authors, other authors are listed alphabetically.

During the preparation of this work, the author(s) used generative AI tools to assist with creating bibtex code for references that were manually listed. After using this tool, the author(s) reviewed and edited the content as necessary and take(s) full responsibility for the accuracy and integrity of the final content of the article.

Competing interests

The authors declare no competing interests.

References

1. Ding, Y., Udompanyawit, C., Zhang, Y. & He, B. EEG-based brain-computer interface enables real-time robotic hand control at individual finger level. *Nat. Commun.* **16**, 4521, DOI: [10.1038/s41467-025-61064-x](https://doi.org/10.1038/s41467-025-61064-x) (2025).

2. Lee, S., Ghajari, S., Sadeghi, S. *et al.* A subnanolitre tetherless optoelectronic microsystem for chronic neural recording in awake mice. *Nat. Electron.* **8**, 830–842, DOI: [10.1038/s41928-025-01484-1](https://doi.org/10.1038/s41928-025-01484-1) (2025).
3. Zang, J. *et al.* A novel wearable device integrating ECG and PCG for cardiac health monitoring. *Microsystems & Nanoeng.* **11**, 7, DOI: [10.1038/s41378-024-00858-3](https://doi.org/10.1038/s41378-024-00858-3) (2025).
4. Kaluza, M., Janik, S., Lubinski, T. *et al.* Clinical validation of noninvasive blood glucose measurements by midinfrared spectroscopy. *Commun. Medicine* **5**, 186, DOI: [10.1038/s43856-025-01241-7](https://doi.org/10.1038/s43856-025-01241-7) (2025).
5. Kaifosh, P. & Reardon, T. R. A generic non-invasive neuromotor interface for human-computer interaction. *Nature* **645**, 702–711 (2025).
6. Slutzky, M. W., Vansteensel, M. J., Herff, C. & Gaunt, R. A. A brain–computer interface working definition. *Nat. Biomed. Eng.* **9**, 512–515, DOI: [10.1038/s41551-025-01414-8](https://doi.org/10.1038/s41551-025-01414-8) (2025).
7. Willett, F. R., Avansino, D. T., Hochberg, L. R. *et al.* High-performance brain-to-text communication via handwriting. *Nature* **593**, 249–254, DOI: [10.1038/s41586-021-03506-2](https://doi.org/10.1038/s41586-021-03506-2) (2021).
8. Qi, Y., Zhu, X., Xiong, X. *et al.* Human motor cortex encodes complex handwriting through a sequence of states. *Nat. Hum. Behav.* **9**, 482–493, DOI: [10.1038/s41562-025-02157-x](https://doi.org/10.1038/s41562-025-02157-x) (2025).
9. Wang, Z., Xu, G., Yu, B. *et al.* Cortical representation of multidimensional handwriting movement and its application to brain-computer interfaces. *Nat. Commun.* **17**, 1420, DOI: [10.1038/s41467-026-70536-7](https://doi.org/10.1038/s41467-026-70536-7) (2026).
10. Willett, F. R., Kunz, E. M., Fan, C. *et al.* A high-performance speech neuroprosthesis. *Nature* **620**, 1031–1036, DOI: [10.1038/s41586-023-06377-x](https://doi.org/10.1038/s41586-023-06377-x) (2023).
11. Wairagkar, M., Card, N. S., Singer-Clark, T. *et al.* An instantaneous voice-synthesis neuroprosthesis. *Nature* DOI: [10.1038/s41586-025-09127-3](https://doi.org/10.1038/s41586-025-09127-3) (2025). In Press.
12. Littlejohn, K. T., Cho, C. J., Liu, J. R. *et al.* A streaming brain-to-voice neuroprosthesis to restore naturalistic communication. *Nat. Neurosci.* **28**, 580–592, DOI: [10.1038/s41593-025-01905-6](https://doi.org/10.1038/s41593-025-01905-6) (2025).
13. Jehi, L., Jette, N., Kwon, C.-S. *et al.* Timing of referral to evaluate for epilepsy surgery: Expert consensus recommendations from the surgical therapies commission of the international league against epilepsy. *Epilepsia* **63**, 2491–2506, DOI: [10.1111/epi.17350](https://doi.org/10.1111/epi.17350) (2022).
14. Quraishi, I. H. & Hirsch, L. J. Long-term ambulatory intracranial eeg. *Stereotact. functional neurosurgery* **103**, 403–414 (2025).
15. Steinmetz, N. A. *et al.* Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science* **372**, eabf4588 (2021).
16. Stevenson, I. H. & Kording, K. P. How advances in neural recording affect data analysis. *Nat. Neurosci.* **14**, 139–142, DOI: [10.1038/nn.2731](https://doi.org/10.1038/nn.2731) (2011).
17. Tchoe, Y., Bourhis, A. M., Cleary, D. R. *et al.* Human brain mapping with multithousand-channel ptnrgrids resolves spatiotemporal dynamics. *Sci. Transl. Medicine* **14**, eabj1441, DOI: [10.1126/scitranslmed.abj1441](https://doi.org/10.1126/scitranslmed.abj1441) (2022).
18. Chatterjee, B., Nath, M., Kumar K, G. *et al.* Biphasic quasistatic brain communication for energy-efficient wireless neural implants. *Nat. Electron.* **6**, 694–705, DOI: [10.1038/s41928-023-01000-3](https://doi.org/10.1038/s41928-023-01000-3) (2023).
19. Sarkar, S., Huang, Q., Antal, S. *et al.* Body-resonance: transmission line-like wireless links enabling high-speed wearable communication. *Commun. Eng.* **4**, 21, DOI: [10.1038/s44172-025-00053-z](https://doi.org/10.1038/s44172-025-00053-z) (2025).
20. Liu, S.-C., Delbruck, T., Indiveri, G., Whatley, A. & Douglas, R. *Event-Based Neuromorphic Systems* (John Wiley & Sons, Chichester, UK, 2015).
21. Donati, E. & Valle, G. Neuromorphic hardware for somatosensory neuroprostheses. *Nat. Commun.* **15**, 556, DOI: [10.1038/s41467-024-44723-3](https://doi.org/10.1038/s41467-024-44723-3) (2024).
22. Bartolozzi, C., Indiveri, G. & Donati, E. Embodied neuromorphic intelligence. *Nat. Commun.* **13**, 1025, DOI: [10.1038/s41467-022-28487-2](https://doi.org/10.1038/s41467-022-28487-2) (2022).
23. Finateu, T. *et al.* 5.10 A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with 4.86µm Pixels, 1.066GEPS Readout, Programmable Event-Rate Controller and Compressive Data-Formatting Pipeline. In *2020 IEEE International Solid-State Circuits Conference - (ISSCC)*, 112–114, DOI: [10.1109/ISSCC19947.2020.9063149](https://doi.org/10.1109/ISSCC19947.2020.9063149) (2020).
24. Yang, M., Chien, C.-H., Delbruck, T. & Liu, S.-C. A 0.5 V 55 µW 64× 2 Channel Binaural Silicon Cochlea for Event-Driven Stereo-Audio Sensing. *IEEE J. Solid-State Circuits* **51**, 2554–2569 (2016).

25. Pehle, C., Billaudelle, S., Cramer, B. *et al.* The brainscales-2 accelerated neuromorphic system with hybrid plasticity (2022).
26. Brink, S., Nease, S., Hasler, P. *et al.* A learning-enabled neuron array ic based upon transistor channel models of biological phenomena. *IEEE Transactions on Biomed. Circuits Syst.* **7**, 71–81 (2012).
27. Davies, M. *et al.* Loihi: A neuromorphic manycore processor with on-chip learning. *Ieee micro* **38**, 82–99 (2018).
28. Cartiglia, M. *et al.* A 4096 channel event-based multielectrode array with asynchronous outputs compatible with neuromorphic processors. *Nat. Commun.* **15**, 7163, DOI: [10.1038/s41467-024-50783-2](https://doi.org/10.1038/s41467-024-50783-2) (2024).
29. Bartels, J. *et al.* Event driven neural network on a mixed signal neuromorphic processor for eeg based epileptic seizure detection. *Sci. Reports* **15**, 15965, DOI: [10.1038/s41598-025-99272-6](https://doi.org/10.1038/s41598-025-99272-6) (2025).
30. Costa, F. *et al.* Robust compression and detection of epileptiform patterns in ECoG using a real-time spiking neural network hardware framework. *Nat. Commun.* **15**, 3275, DOI: [10.1038/s41467-024-47495-y](https://doi.org/10.1038/s41467-024-47495-y) (2024).
31. De Luca, C., Tincani, M., Indiveri, G. & Donati, E. A neuromorphic multi-scale approach for real-time heart rate and state detection. *npj Unconv. Comput.* **2**, 6 (2025).
32. Zanghieri, M. *et al.* Event-based estimation of hand forces from high-density surface emg on a parallel ultralow-power microcontroller. *IEEE Sensors J.* **25**, 7771–7780 (2024).
33. Narayanan, S., Ghosh, S. & Indiveri, G. A 32-channel event-based bio-signal analog front-end with adaptive delta and pulse frequency encoding. In *arXiv preprint arXiv:2607.12901*, DOI: [10.48550/arXiv.2607.12901](https://doi.org/10.48550/arXiv.2607.12901) (2026).
34. Corradi, F., Bontrager, D. & Indiveri, G. Toward neuromorphic intelligent brain-machine interfaces: An event-based neural recording and processing system. In *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*, 584–587 (IEEE, 2014).
35. Graca, R., Zhou, S., McReynolds, B. & Delbruck, T. Scidvs: A scientific event camera with 1.7% temporal contrast sensitivity at 0.7 lux. In *2024 IEEE European Solid-State Electronics Research Conference (ESSERC)*, 205–208 (IEEE, 2024).
36. Gallego, G. *et al.* Event-based vision: A survey. *IEEE transactions on pattern analysis machine intelligence* **44**, 154–180 (2020).
37. Lichtsteiner, P. & Delbruck, T. A 64x64 aer logarithmic temporal derivative silicon retina. In *Research in Microelectronics and Electronics, 2005 PhD*, vol. 2, 202–205, DOI: [10.1109/RME.2005.1542972](https://doi.org/10.1109/RME.2005.1542972) (2005).
38. Montemurro, M. A., Rasch, M. J., Murayama, Y., Logothetis, N. K. & Panzeri, S. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr. Biol.* **18**, 375–380, DOI: [10.1016/j.cub.2008.02.023](https://doi.org/10.1016/j.cub.2008.02.023) (2008).
39. Erturk, E., Hashemi, S. & Shanechi, M. Cross-modal representational knowledge distillation for enhanced spike-informed lfp modeling. In *Advances in Neural Information Processing Systems*, vol. 38 (Curran Associates, Inc., 2025).
40. Izhikevich, E. M. Bursting. *Scholarpedia* **1**, 1300, DOI: [10.4249/scholarpedia.1300](https://doi.org/10.4249/scholarpedia.1300) (2006). Revision #142385.
41. Belitski, A. *et al.* Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J. Neurosci.* **28**, 5696–5709 (2008).
42. Inose, H., Aoki, T. & Watanabe, K. Asynchronous delta-modulation system. *Electron. Lett.* **3**, 95–96 (1966).
43. Hawkes, T. & Simonpieri, P. Signal coding using asynchronous delta modulation. *IEEE Transactions on Commun.* **22**, 729–731 (1974).
44. Lapique, L. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *J. de Physiol. et de Pathol. Générale* **9**, 620–635 (1907).
45. Thorp, C. K. & Steinmetz, P. N. Interference and noise in human intracranial microwire recordings. *IEEE Transactions on Biomed. Eng.* **56**, 30–36 (2008).
46. Keshtkaran, M. R. & Yang, Z. A fast, robust algorithm for power line interference cancellation in neural recording. *J. neural engineering* **11**, 026017 (2014).
47. Shimoda, K., Nagasaka, Y., Chao, Z. C. & Fujii, N. Decoding continuous three-dimensional hand trajectories from epidural electrocorticographic signals in japanese macaques. *J. neural engineering* **9**, 036015 (2012).
48. Quiroga, R. Q., Nadasdy, Z. & Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural computation* **16**, 1661–1687 (2004).

49. Blau, Y. & Michaeli, T. Rethinking lossy compression: The rate-distortion-perception tradeoff. In *Proceedings of the 36th International Conference on Machine Learning*, vol. 97 of *Proceedings of Machine Learning Research*, 675–685, DOI: [10.48550/arXiv.1901.07821](https://doi.org/10.48550/arXiv.1901.07821) (PMLR, 2019).
50. Izhikevich, E. M. Neural excitability, spiking, and bursting. *Int. J. Bifurc. Chaos* **10**, 1171–1266, DOI: [10.1142/s0218127400000840](https://doi.org/10.1142/s0218127400000840) (2000).
51. Tegnigkkeit, F., Schwarz, D. & Puil, E. Modulation of bursts and high-threshold calcium spikes in neurons of rat auditory thalamus. *Neuroscience* **83**, 1063–1073, DOI: [10.1016/s0306-4522\(97\)00458-2](https://doi.org/10.1016/s0306-4522(97)00458-2) (1998).
52. Chacron, M. J., Longtin, A. & Maler, L. To burst or not to burst? *J. Comput. Neurosci.* **17**, 127–136, DOI: [10.1023/b:jcms.0000037677.58916.6b](https://doi.org/10.1023/b:jcms.0000037677.58916.6b) (2004).
53. Su, H., Alroy, G., Kirson, E. D. & Yaari, Y. Extracellular calcium modulates persistent sodium current-dependent burst-firing in hippocampal pyramidal neurons. *The J. Neurosci.* **21**, 4173–4182, DOI: [10.1523/jneurosci.21-12-04173.2001](https://doi.org/10.1523/jneurosci.21-12-04173.2001) (2001).
54. Graves, A. R. *et al.* Hippocampal pyramidal neurons comprise two distinct cell types that are countermodulated by metabotropic receptors. *Neuron* **76**, 776–789, DOI: [10.1016/j.neuron.2012.09.036](https://doi.org/10.1016/j.neuron.2012.09.036) (2012).
55. Hwang, C. *et al.* Event-based neural spike detection using spiking neural networks for neuromorphic ibmi systems. In *2025 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5 (IEEE, 2025).
56. He, Y. *et al.* An event-based neural compressive telemetry with $> 11\times$ loss-less data reduction for high-bandwidth intracortical brain computer interfaces. *IEEE transactions on biomedical circuits systems* **18**, 1100–1111 (2024).
57. Swec-ethz ieeg database. webpage <http://ieeg-swez.ethz.ch/>. recorded and analysed at Sleep Wake Epilepsy Center (SWEC), University Department of Neurology, Inselspital Bern and at Integrated Systems Laboratory, ETH Zurich.
58. Burrello, A., Schindler, K., Benini, L. & Rahimi, A. One-shot learning for iEEG seizure detection using end-to-end binary operations: Local binary patterns with hyperdimensional computing. In *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 1–4 (IEEE, 2018).
59. Mormann, F., Andrzejak, R. G., Elger, C. E. & Lehnertz, K. Seizure prediction: the long and winding road. *Brain* **130**, 314–333 (2007).
60. Mormann, F. & Andrzejak, R. G. Seizure prediction: making mileage on the long and winding road. *Brain* **139**, 1625–1627 (2016).
61. Andrzejak, R. G. *et al.* Seizure forecasting: Where do we stand? *Epilepsia* **64**, S62–S71 (2023).
62. He, Y. *et al.* An Event-Based Neural Compressive Telemetry With $> 11\times$ Loss-Less Data Reduction for High-Bandwidth Intracortical Brain Computer Interfaces. *IEEE Transactions on Biomed. Circuits Syst.* **18**, 1100–1111 (2024).
63. Biederman, W. *et al.* A 4.78 mm² Fully-Integrated Neuromodulation SoC Combining 64 Acquisition Channels With Digital Compression and Simultaneous Dual Stimulation. *IEEE J. Solid-State Circuits* **50**, 1038–1047 (2015).
64. Shaeri, M. A. *et al.* 33.3 MiBmi: A 192/512-Channel 2.46mm² Miniaturized Brain-Machine Interface Chipset Enabling 31-Class Brain-to-Text Conversion Through Distinctive Neural Codes. In *2024 IEEE International Solid-State Circuits Conference (ISSCC)*, 546–548 (IEEE, San Francisco, CA, USA, 2024).
65. Tsai, D., Sawyer, D., Bradd, A., Yuste, R. & Shepard, K. L. A very large-scale microelectrode array for cellular-resolution electrophysiology. *Nat. Commun.* **8**, 1802 (2017).
66. Lopez, C. M. *et al.* A Multimodal CMOS MEA for High-Throughput Intracellular Action Potential Measurements and Impedance Spectroscopy in Drug-Screening Applications. *IEEE J. Solid-State Circuits* **53**, 3076–3086 (2018).
67. Dragas, J. *et al.* In Vitro Multi-Functional Microelectrode Array Featuring 59 760 Electrodes, 2048 Electrophysiology Channels, Stimulation, Impedance Measurement, and Neurotransmitter Detection Channels. *IEEE J. Solid-State Circuits* **52**, 1576–1590 (2017).
68. Pehle, C. *et al.* The brainscales-2 accelerated neuromorphic system with hybrid plasticity. *Front. Neurosci.* **16**, 795876 (2022).
69. Qiao, N. *et al.* A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128k synapses. *Front. neuroscience* **9**, 141 (2015).
70. Brink, S. *et al.* A learning-enabled neuron array ic based upon transistor channel models of biological phenomena. *IEEE Transactions on Biomed. Circuits Syst.* **7**, 71–81 (2012).

71. Müller, E. *et al.* A scalable approach to modeling on accelerated neuromorphic hardware. *Front. neuroscience* **16**, 884128 (2022).
72. Heck, C. N. & *et. al.* Two-year seizure reduction in adults with medically intractable partial onset epilepsy treated with responsive neurostimulation: Final results of the RNS System Pivotal trial. *Epilepsia* **55**, 432–441, DOI: [10.1111/epi.12534](https://doi.org/10.1111/epi.12534) (2014).
73. Duun-Henriksen, J. & *et. al.* A new era in electroencephalographic monitoring? subscalp devices for ultra-long-term recordings. *Epilepsia* **61**, 1805–1817, DOI: [10.1111/epi.16630](https://doi.org/10.1111/epi.16630) (2020).
74. Halliday, A. J. & *et. al.* The UMPIRE study: A first-in-human multicenter trial of bilateral subscalp monitoring for epileptic seizure detection. *Epilepsia* **66**, 3426–3439, DOI: [10.1111/epi.18458](https://doi.org/10.1111/epi.18458) (2025).
75. Pavan, S., Schreier, R. & Temes, G. C. *Understanding Delta-Sigma Data Converters* (John Wiley & Sons, Hoboken, NJ, USA, 2017), 2nd edn.
76. Mohan, V., Tay, W. & Basu, A. Towards neuromorphic compression based neural sensing for next-generation wireless implantable brain machine interface. *Neuromorphic Comput. Eng.* **5**, 014004 (2025).
77. Safa, A. *et al.* Exploring information-theoretic criteria to accelerate the tuning of neuromorphic level-crossing adcs. In *Proceedings of the 2023 Annual Neuro-Inspired Computational Elements Conference*, 63–70 (2023).
78. Burnham, K. P. & Anderson, D. R. *Model selection and multimodel inference: a practical information-theoretic approach* (Springer, 2002).
79. Fang, W. *et al.* Incorporating learnable membrane time constant to enhance learning of spiking neural networks. In *Proceedings of the IEEE/CVF international conference on computer vision*, 2661–2671 (2021).
80. Zhang, Z. & Constandinou, T. G. Adaptive spike detection and hardware optimization towards autonomous, high-channel-count bmis. *J. Neurosci. Methods* **354**, 109103 (2021).
81. Yin, B., Corradi, F. & Bohtë, S. M. Accurate and efficient time-domain classification with adaptive spiking recurrent neural networks. *Nat. Mach. Intell.* **3**, 905–913 (2021).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementary.pdf](#)