



Sub-millisecond closed-loop feedback stimulation between arbitrary sets of individual neurons

Jan Müller*, Douglas J. Bakkum and Andreas Hierlemann

Bio Engineering Laboratory, ETH Zürich, Basel, Switzerland

Edited by:

Steve M. Potter, Georgia Institute of Technology, USA

Reviewed by:

Antonio Novellino, ETT s.r.l., Italy
Jürg Streit, University of Bern, Switzerland

*Correspondence:

Jan Müller, Bio Engineering Laboratory, ETH Zürich, Basel, Switzerland.
e-mail: 217534@gmail.com

We present a system to artificially correlate the spike timing between sets of arbitrary neurons that were interfaced to a complementary metal–oxide–semiconductor (CMOS) high-density microelectrode array (MEA). The system features a novel reprogrammable and flexible event engine unit to detect arbitrary spatio-temporal patterns of recorded action potentials and is capable of delivering sub-millisecond closed-loop feedback of electrical stimulation upon trigger events in real-time. The relative timing between action potentials of individual neurons as well as the temporal pattern among multiple neurons, or neuronal assemblies, is considered an important factor governing memory and learning in the brain. Artificially changing timings between arbitrary sets of spiking neurons with our system could provide a “knob” to tune information processing in the network.

Keywords: closed-loop, high-density microelectrode array, STDP, acausal stimulation, LTD, sub-millisecond

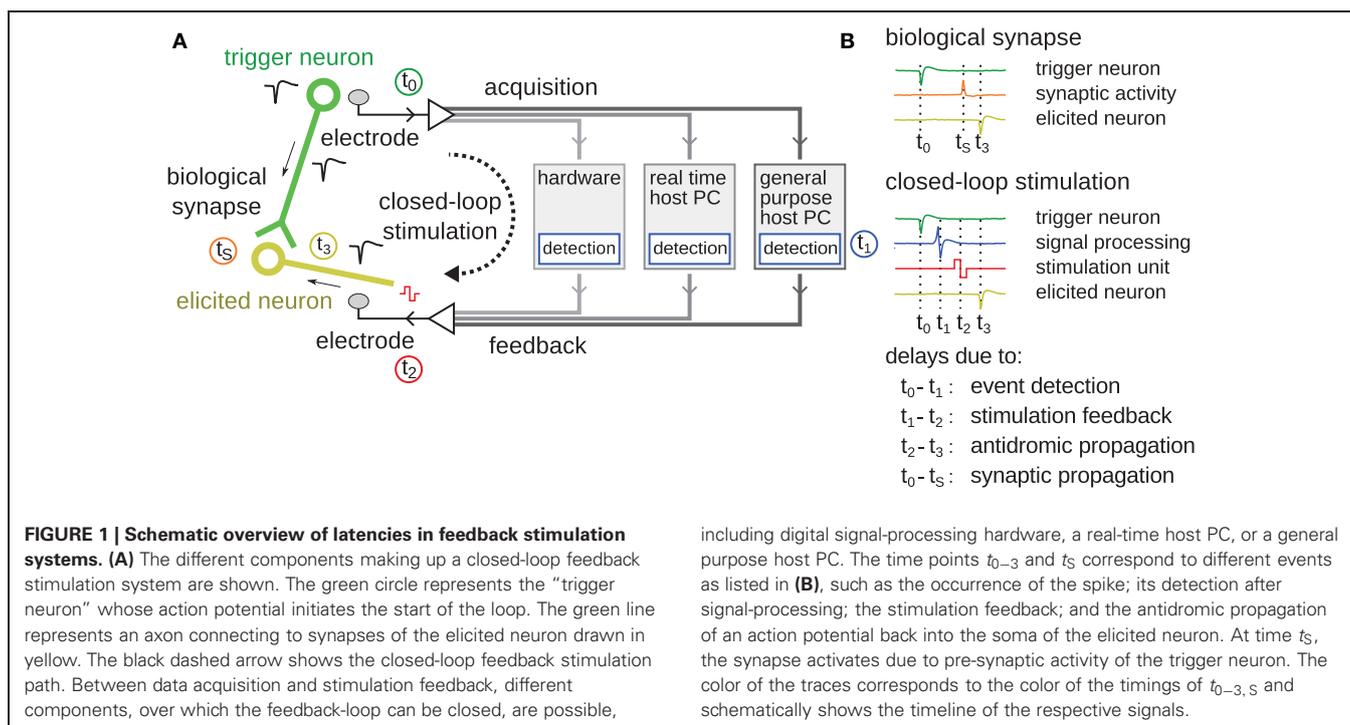
INTRODUCTION

Different theories describing learning and memory in the brain have been developed, and converging evidence shows that the precise activity timing of individual or groups of neurons may play a paramount role in plasticity of neuronal circuits. The well-known spike timing dependent plasticity (STDP) rule states that if two synaptically connected neurons fire within tens of milliseconds of each other, the connectivity strength of the involved synapses gets potentiated or depressed depending on the firing order. In pioneering studies, STDP rules were discovered (Markram et al., 1997) and further characterized (Bi and Poo, 1998; Song et al., 2000) by observing the effect of correlated firing of two neurons either artificially induced by stimulating a pre- and a post-synaptic neuron with two patch-clamps or by applying trains of paired-pulse stimuli to one neuron in the network (Bi and Poo, 1999). Furthermore, computation in a network is likely due not only to the relative timing of two individual neurons but also to the correlated activity of different neurons forming an associated group, i.e., assembly (Chang et al., 2000; Izhikevich, 2006). In this vein, different studies reported the existence of precise time-locked activity patterns of multiple neurons, both *in vivo* and *in vitro* (Abeles and Gerstein, 1988; Bienenstock, 1995; Ikegaya et al., 2004; Rolston et al., 2007). Having a system to generate feedback stimulation quickly and accurately to interact with such activity patterns would expand such studies beyond finding rules governing the plasticity between two cells toward finding rules governing the spatio-temporal dynamics of whole networks or assemblies (Froemke and Dan, 2002; Izhikevich et al., 2004).

In recent years, different systems to artificially control such feedback stimulation in a closed-loop manner, and thus study neuronal plasticity, have been developed for both *in vivo* (Jackson et al., 2006b; Bontorin et al., 2007; Venkatraman et al., 2009) and *in vitro* applications (Bontorin et al., 2007; Hafizovic et al., 2007; Novellino et al., 2007; Rolston et al., 2010; Zrenner et al., 2010; Wallach et al., 2011). In turn, activity-dependent feedback

stimulation was shown to modify the functional connectivity of neuronal networks, both *in vivo* and *in vitro*, as done by reprogramming the motor output of freely behaving primates (Jackson et al., 2006a), changing the functional connectivity in rat forelimb sensorimotor cortex (Rebesco et al., 2010), or shaping *in vitro* neocortical networks into predefined activity states (Bakkum et al., 2008b). *In vivo* systems usually record from needles inserted into a certain location of the brain and subsequently stimulate the same or another site upon the detection of activity. These systems usually comprise the implanted needles, a head stage to amplify the signals, and some means to transmit the acquired signals to a PC. In the case of closed-loop feedback stimulation, these systems usually feature a dedicated very-large-scale-integrated application-specific circuit (VLSI ASIC) (Chen et al., 2009; Rizk et al., 2009; Lee et al., 2010; Azin et al., 2011), or use a general-purpose microcontroller to achieve the respective goals (Mavoori et al., 2005; Zanos et al., 2011). Most *in vitro* systems, on the other hand, use a data acquisition card (DAQ) to sample data for analysis on a PC; feedback stimulation is typically returned through a DAQ card as well.

In order to accurately control the timing of feedback stimulation loops within the timescales relevant for STDP to occur, the delays introduced by a system must be understood. A generic description is given in **Figure 1**. Different system implementations will have different sources for and values of delays. Signal-processing algorithms introduce an inherent delay in the processing itself. Systems, which rely on general-purpose computers, might introduce latencies and jitter through the presence of data buffers, interrupts, shared resources, or user interactions, etc. In **Figure 1**, the time points t_{0-3} and t_5 specify the occurrence of important events. At $t_0 = 0$, the trigger neuron emits an action potential, which is recorded by the acquisition system. After entering the signal-processing stages, it is ready to be detected as a spike event at time t_1 . From there, the system emits a stimulation pulse hitting the electrode at time t_2 . Conventionally,



the loop is considered “closed” at this point. The stimulation pulse evokes neuronal activity, frequently activating nearby axons (Bakkum et al., 2008a) whose signals propagate antidromically toward the soma until eliciting an action potential at time t_3 . In the case depicted in **Figure 1**, where the trigger neuron is synaptically connected to the elicited neuron, an additional biological time, t_s , denotes the duration of an action potential propagation through the axon of the trigger neuron until synaptic activation of the elicited neuron. In case where $t_0 - t_1 - t_2$ is faster than $t_0 - t_s$, that is when the signal propagates faster through the artificial feedback-loop than down the axon toward the biological synapse, acausal stimulation, and thus the introduction of long-term depression (LTD) according to the STDP rule, is possible.

In order to apply closed-loop stimulation feedback precise and fast enough to study plasticity at the timescales of STDP or acausal stimulation, and flexible enough to interact with cell assemblies, we developed a field-programmable gate array (FPGA)-based system, interfaced with a complementary metal–oxide–semiconductor high-density microelectrode array (CMOS-MEA). The CMOS-MEA features a total of 126 read-out and 42 stimulation channels, which can be connected to an almost arbitrary subset of $11,011.5 \times 7 \mu\text{m}^2$ electrodes, arranged in a $2 \times 1.75 \text{ mm}^2$ array. The feedback stimulation loop is closed around the CMOS-MEA using an FPGA that performs signal-processing, such as spike-detection and feedback generation. The system functionality was verified using cultured networks of cortical neurons and glia. The minimum programmable latency of the closed-loop stimulation feedback ($t_0 - t_1 - t_2$) was $400 \mu\text{s}$ with jitter below $50 \mu\text{s}$, suitable to induce STDP. This is faster than many axonal propagation delays ($t_0 - t_s$), rendering it possible to

including digital signal-processing hardware, a real-time host PC, or a general purpose host PC. The time points t_{0-3} and t_s correspond to different events as listed in (B), such as the occurrence of the spike; its detection after signal-processing; the stimulation feedback; and the antidromic propagation of an action potential back into the soma of the elicited neuron. At time t_s , the synapse activates due to pre-synaptic activity of the trigger neuron. The color of the traces corresponds to the color of the timings of t_{0-3} and schematically shows the timeline of the respective signals.

conduct acausal stimulation experiments. An “event engine” was designed and implemented to trigger feedback stimulation at the occurrence of activity patterns, such as those described in Ikegaya et al. (2004) and Rolston et al. (2007). Patterns could be of almost arbitrary length and could consist of up to 1000’s of individual elements, only limited by the available resources of the FPGA. Configurations for the event engine could be (re)loaded within milliseconds. Unique to this system is the possibility to enable low-latency, high-throughput, STDP-like experiments as well as acausal stimulations across many individual neurons, or neuronal assemblies in parallel through the simultaneous application of many feedback stimulation loops. To infer changes in synaptic strengths, correlations between putative mono-synaptically connected neurons (Fujisawa et al., 2008) can be monitored using extracellular spikes. In the future, high-throughput STDP experiments will be possible by adding a patch electrode to the system in order to monitor changes in intracellular post-synaptic currents.

METHODS

SYSTEM ARCHITECTURE

The main design goals were to implement (1) multiple feedback stimulation loops (2) to match arbitrary spike patterns with (3) short latencies ($<1 \text{ ms}$) and (4) high accuracy ($<50 \mu\text{s}$) (5) while still recording from all available 126 channels. A main component of the presented system is an FPGA, used to hijack signals traveling between the analog-to-digital converter on the CMOS device and the host PC. Due to the inherent parallel nature of FPGAs, signal-processing and feedback generation using data from additional recording channels can be done without introducing additional delays or jitter.

The system consists of three main parts as shown in **Figure 2**. The first is a high-density CMOS-MEA device featuring on chip signal-conditioning, stimulation, and analog-to-digital conversion (ADC) units (Frey et al., 2010), described in more detail in the next section. It is plugged into a custom printed circuit board (PCB) that provides reference voltages and clock signals. The digital data as provided by the CMOS-MEA are transmitted through a low-voltage differential link to reduce sensitivity to electromagnetic interferences as caused, for example, by a nearby incubator. The second part is an FPGA, which reads in the differential signals and subsequently performs signal-processing, spike-detection, and feedback stimulation, as well as compression and framing of the data to be sent via TCP/IP over Ethernet to a host PC, the third main part. On the host PC, further data analysis can be performed online or offline. It is also used to program and control the CMOS-MEA device during experimentation with different settings, like amplifier gain or electrode-to-amplifier routing, in order to be adopted for use in different experimental sessions.

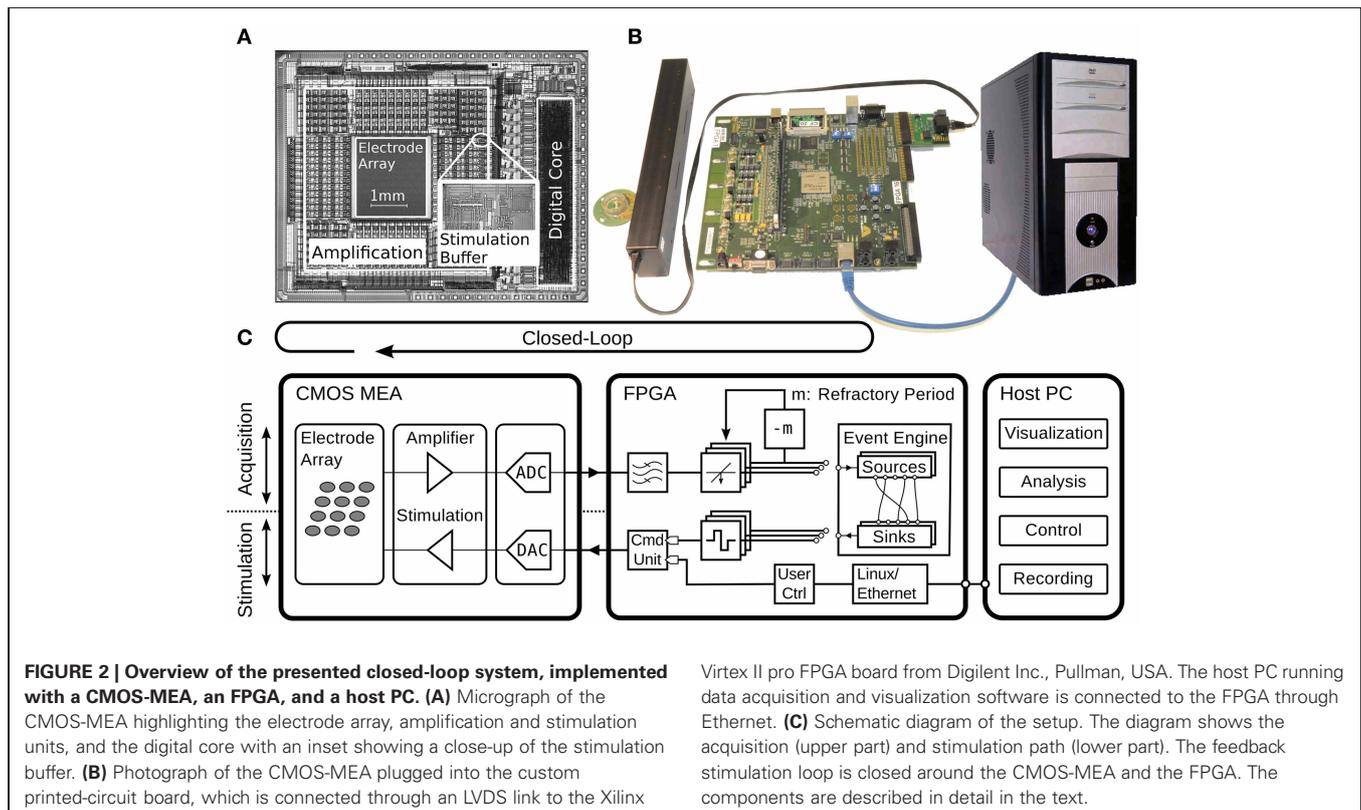
CMOS DEVICE

The CMOS-MEA includes 126 readout channels with programmable amplification (0 dB to 80 dB), on chip ADCs sampling at 20 kHz, and stimulation capabilities (see below). It features a sensor area of $2 \times 1.75 \text{ mm}^2$ with a total of 11,011 electrodes, each with a size of $5 \times 7 \mu\text{m}^2$ and a pitch of $18 \mu\text{m}$. Beneath the electrodes resides a sophisticated analog-switching matrix to connect an almost arbitrary subset of the 11,011 electrodes to the 126 readout channels. The readout electronics were placed

outside of the sensor array, instead of directly below the electrodes as done in active-pixel sensor devices (APS) (Berdoncini et al., 2009), to provide space for larger circuitry elements that produce less noise. This scheme also allows for reducing the pitch of the electrodes below the spatial requirements of the readout electronics. See Frey et al. (2010) for more details.

FPGA

A reprogrammable Virtex II pro FPGA (Xilinx Inc., San Jose, USA) was used as an intermediate signal-processing device between the CMOS-MEA and the host PC to perform real-time signal-processing, decision-making and feedback generation. The FPGA acquires digital data coming from the differential link and forwards it to a PC over Ethernet. The Virtex II pro features an embedded PowerPC microprocessor running at 300 MHz that operates a Linux kernel with a Busybox operating system. The TCP/IP stack of the Linux kernel handles the network communication and data transfer. As the embedded PowerPC microprocessor is relatively slow, compared to modern CPUs, this provides a bottleneck for fast data transmission. We measured the latency between the TCP/IP stack of the FPGA and the host PC to be $83 \pm 21 \text{ ms}$ (mean \pm SD, $N = 308$) at full-frame data transmission, which is larger than the STDP window of up to tens of milliseconds. One solution to this problem might be to stop streaming of the full data readout, while performing a closed-loop experiment and to only route out the data channels strictly needed for the closed-loop feedback stimulation. This would free some of the bandwidth of the Ethernet link and make it available for faster feedback stimulation. Crucially, however,



we would lose the possibility to simultaneously monitor neural activity elsewhere in the cultured network by applying such a paradigm. Another option might be to bypass the Ethernet link by streaming the data directly to a DAQ card, attached to the host PC, and to send stimulation information back through a second link to the FPGA. All these methods are less practical than using the universal TCP/IP connection, which plugs into almost every kind of host PC and does not require additional hardware. An attractive alternative for achieving low latencies was to implement all needed signal-processing and feedback generation directly on the FPGA. The next paragraphs highlight the different building blocks needed to implement such a scheme. Although the FPGA can be reprogrammed at will, this is time-consuming and error prone and, therefore, not suitable during an experimental session. To accommodate reprogramming, a more flexible, module-based design was developed in VHDL and programmed into the FPGA logic together with a software interface to quickly reconfigure the connectivity of the individual modules (see “Event Engine”).

SPIKE-DETECTION

One such signal-processing building block is spike-detection, which extracts spiking events from the raw voltage traces, recorded at the electrodes. Spike-detection is implemented as a threshold crossing. The signals are first digitally band-pass filtered with a two-tab Butterworth filter (500 Hz–3 kHz) to suppress DC offset components and higher frequency noise; this will emphasize the action potential frequency components. The detection threshold level is user-programmable and typically set around 4.5 times the noise standard deviation. During experimentation, this value can be determined by software running online on the host PC. After an identified spike event, we set a programmable refractory period to 3 ms. After stimulation, detection was disabled for 3 ms as well, to avoid oscillating loops due to feedback stimulation artifacts being falsely classified as spikes.

EVENT ENGINE

To avoid time-consuming reprogramming of the FPGA fabric, a more flexible and modular event-based scheme for feedback generation (Event Engine) was designed and implemented. The event engine consists of small building blocks, called modules, each of which implements a specific simple function. Each module has one or more event sinks as inputs and one event source as an output. By connecting the event sources to the appropriate event sinks, different, almost arbitrary pattern matching, and event handling algorithms can be achieved. **Table 1** summarizes the implemented modules. **Figure 3** shows different basic configurations to achieve defined pattern matching. In **Figure 3A**, the simplest closed-loop configuration is depicted, where the source of a spike-detection module gets connected to the sink of a delay unit and from there to a stimulation function generator. Whenever the source produces an event (i.e., in this case detects a spike), the sink triggers a stimulation pulse after a defined time delay. By means of software, the sources can be connected to sinks dynamically and rapidly within milliseconds while running an experiment such that pattern matching can adapt to ongoing activity in the living culture. One notable property is the lack of time binning. Each spike gets represented as a single pulse with a

temporal resolution set by the sampling frequency, i.e., 20 kHz. As a consequence, certain desired operations might not make sense, as the biological neurons have some inherent variability in when they spike. For example, the user might want to match a pattern, where two neurons spike together (see **Figure 3E**). To achieve this, a SPREADING module “spreads” the spike pulse in time in order to compensate for jitter. This way, the subsequent AND module can generate an output event whenever the two neurons fire together within a specified range of time. As discussed in Ikegaya et al. (2004) and Rolston et al. (2007), 2 ms is suitable for most recurring patterns. Another module can be used to convert the spread-out spike pulse back into a single one-shot event, which then can be used, for example, to trigger the stimulation unit only once per spread-out pulse. The particular selection of implemented modules (as listed in **Table 1**) represents a minimal set, which, if combined in the appropriate way, allows for matching different kinds of events, such as specific spatio-temporal activity patterns, time sequences, network bursts, local bursts, etc. In order to keep the event engine as flexible as possible and adaptable to different, possibly unforeseen pattern matching sequences, the implementation of a minimal set of small building blocks has been chosen over the approach, where each envisioned pattern would require a single, but more complex, and less flexible building block. Thus, available modules can be combined together in almost infinite different ways, limited only by the available FPGA memory that keeps track of all source-sink associations.

STIMULATION/FUNCTION GENERATOR

The CMOS-MEA has 42 on-chip integrated stimulation units, which are driven by two 10bit DACs. On the FPGA is a function generator implemented to achieve arbitrary stimulation waveforms. A defined waveform has to be programmed at the start of the experiment. We used biphasic, first positive then negative voltage pulses of 200 μ s duration per phase and ± 300 or 400 mV amplitude. The stimulation buffers can be chosen to operate in voltage- or current mode (Livi et al., 2010). Whenever the event engine outputs an event, the appropriate stimulation buffer, located on the CMOS-MEA, gets connected, and the function generator starts its operation. Stimulation artifacts on the readout channels could result in falsely detected spikes and cause a reverberation problem for low-latency feedback-loops. Therefore, spike-detection is blanked during a time period of a few milliseconds after stimulation onset.

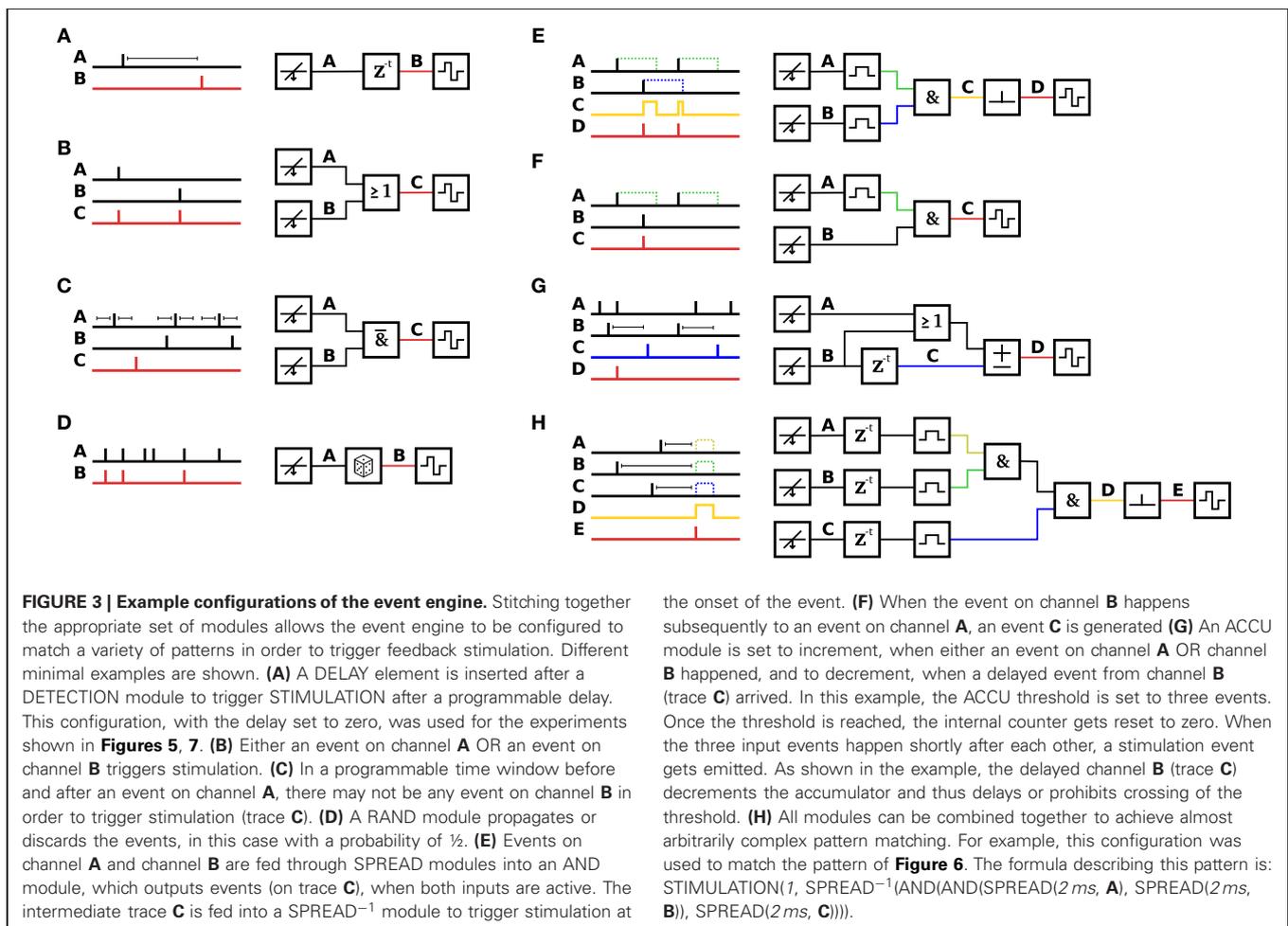
CULTURES

The performance of the closed-loop system was tested with cortical neurons and glia grown over the CMOS-MEA. Animal handling protocols were approved by the Basel-Stadt Veterinary office according to Swiss federal laws on animal welfare. Briefly, a time-pregnant rat was anesthetized using isoflurane, then decapitated to gain E18 embryos. Cortices were extracted from the embryos and dissociated enzymatically in trypsin (Invitrogen) followed by mechanical trituration. A layer of laminin (Sigma) over a layer of poly(ethyleneimine) (Sigma) was used to adhere between 20 and 40 k cells. Plating media consisted of 850 μ L of Neurobasal, supplemented with 10% horse serum (HyClone), 0.5 mM GlutaMAX (Invitrogen), and 2% B27 (Invitrogen). After

Table 1 | A minimal set of modules making up the event engine.

	DELAY(<i>t</i> , A)	Delays the event A by a defined amount of time <i>t</i> .
	AND(A , B)	Emits an event, when both of the two input events, A and B occurred simultaneously.
	OR(A , B)	Emits an event, when either of the two input events A or B occurred.
	INH(<i>t</i> , A , B)	Emits an event, when an event on A , however, no event on B occurred in a defined time window, <i>t</i> , in order to create inhibitory feedback-loops.
	RAND(<i>p</i> , A)	Propagates the event A to the output or drops it after a Bernoulli-distributed pseudo-random variable with a definable probability, <i>p</i> .
	ACCU(<i>n</i> , A , B)	Increments (event A) or decrements (event B) an internal accumulator and emits an event after a definable threshold, <i>n</i> , has been reached, after which it is reset to zero.
	SPREAD(<i>t</i> , A)	Spreads the event A in time for a defined time, <i>t</i> .
	SPREAD ⁻¹ (A)	Converts the onset of a spread-out event A back into a single event.
	DETECTION(<i>c</i>)	Emits an event, when the specified channel, <i>c</i> , detected a spike.
	STIMULATION(<i>c</i> , A)	Generates a stimulation pulse on the specified channel, <i>c</i> , whenever input event A happened.
	START	Single pulse after system start-up, which can be used to start repetitive stimulation protocols.

Configurable parameters are represented in italics (*t*, *p*, *n*, *c*), and input events are denoted in bold letters (**A**, **B**).



24 h, the plating media was changed to growth media: 850 μL of DMEM (Invitrogen), supplemented with 10% horse serum, 0.5 mM GlutaMAX, and 1 mM sodium pyruvate (Invitrogen). Cultures matured for 3–4 weeks prior to experimentation, and experiments were conducted inside an incubator to control environmental conditions (34.5°C and 5% CO_2). For further details see Hales et al. (2010).

EVALUATION AND RESULTS

This section begins with data characterizing the suitability of our setup to perform closed-loop feedback stimulation experiments, using cultures of cortical neurons and glia for validation. First, the process of identifying neurons to be used in closed-loop feedback stimulation will be described. Then the system's loop speed and jitter performance will be quantified. An example event engine was run to provide stimulation feedback, triggered by an activity pattern. Preliminary data and techniques to analyze the consequences of such stimulation on the functional connectivity between neurons will be presented and discussed. Finally, an experimental session to induce LTD through acausal stimulation will be sketched, and its implications discussed. Data in the figures demonstrate proof-of-principle experiments from individual cultures, the setup has, however, been successfully applied to many tens of cultures.

RECORDING/STIMULATION SELECTIVITY

High-density CMOS-MEAs can potentially sample from complete neuronal populations. Due to the high-density (18 μm pitch) of the CMOS electrode array, every neuron lying on the $2 \times 1.75 \text{ mm}^2$ array can be bidirectionally addressed. On the other hand, when stimulating one electrode, a defined subset of neurons is often directly activated in response (Bakkum et al., 2008a). **Figure 4** shows such a scenario. In **Figure 4A**, one electrode, marked with a black cross, was stimulated multiple times, and the evoked activity was recorded during a window of 12 ms after stimulation onset. The median calculated over all voltage traces filters out noise and spontaneously spiking neurons/traces. Reliable activity (usually with a jitter on the order of 100 μs or below) is considered due to an antidromic action potential initiated at the neuron's axon (Lipski, 1981). Since only a subset of 126 out of the 11,011 electrodes can be readout simultaneously, the stimulation sequence was repeated multiple times, each time with a different subset of electrodes, until all electrodes were covered. After recording all sequences, the traces of the individual recordings were aligned in time. To highlight the electrodes that recorded elicited action potentials, the negative peak of the recorded voltage level during 12 ms after stimulation is color-coded and clipped at $-100 \mu\text{V}$. The red circles around the exemplified 11 spots highlight neurons that fired directly elicited action potentials. Their traces are individually shown in **Figure 4B**, demonstrating that the elicited action potentials were reliably and precisely fired after a given time, and only in a few cases (traces 2, 4, 6, 9), activity with different timing occurred. These could stem from a different neuron that happened to sit near the same electrode and/or from action potentials occurring within a coincident network burst.

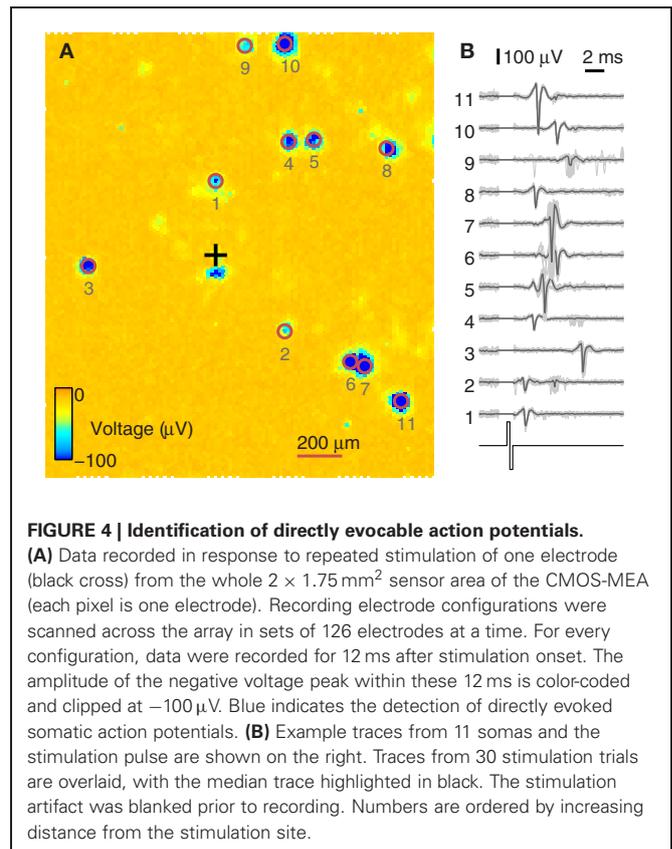


FIGURE 4 | Identification of directly evocable action potentials.

(A) Data recorded in response to repeated stimulation of one electrode (black cross) from the whole $2 \times 1.75 \text{ mm}^2$ sensor area of the CMOS-MEA (each pixel is one electrode). Recording electrode configurations were scanned across the array in sets of 126 electrodes at a time. For every configuration, data were recorded for 12 ms after stimulation onset. The amplitude of the negative voltage peak within these 12 ms is color-coded and clipped at $-100 \mu\text{V}$. Blue indicates the detection of directly evoked somatic action potentials. **(B)** Example traces from 11 somas and the stimulation pulse on the right. Traces from 30 stimulation trials are overlaid, with the median trace highlighted in black. The stimulation artifact was blanked prior to recording. Numbers are ordered by increasing distance from the stimulation site.

As shown, recording and stimulation with the CMOS-MEA feature high spatial resolution and, therefore, are locally very confined. However, the facts that one electrode can detect signals from more than one neuron, and that the stimulation through one electrode can directly evoke action potentials of more than one neuron have to be considered when planning closed-loop feedback stimulation experiments. In this case, the feedback-loop is not closed between two neurons, but includes two sets of neurons.

FEEDBACK LATENCIES

According to the rules of STDP, the timing window to induce long-term potentiation (LTP) at synapses is between less than a few milliseconds and up to tens of milliseconds post-synaptic activation before and after pre-synaptic activity. Thus, even though feedback cycles of 5–10 ms are fast enough to induce LTP, we aimed at reaching cycle-times below 1 ms to enable the system to perform acausal stimulations, as explained in the respective section below.

Figure 5 shows the overlay of 128 traces of the feedback-loop. Here, the event engine was configured to detect events on only one channel and stimulate immediately after detection, i.e., without any further delays in order to test the system performance (cf. **Figure 3A**). The traces are aligned at the onset-time of the stimulation pulse, and time zero is set to be at the negative peak of the spike of the trigger neuron. In red are the traces from the trigger neuron, and in black, the traces from the elicited neuron. The timing between a trigger neuron spike and the onset

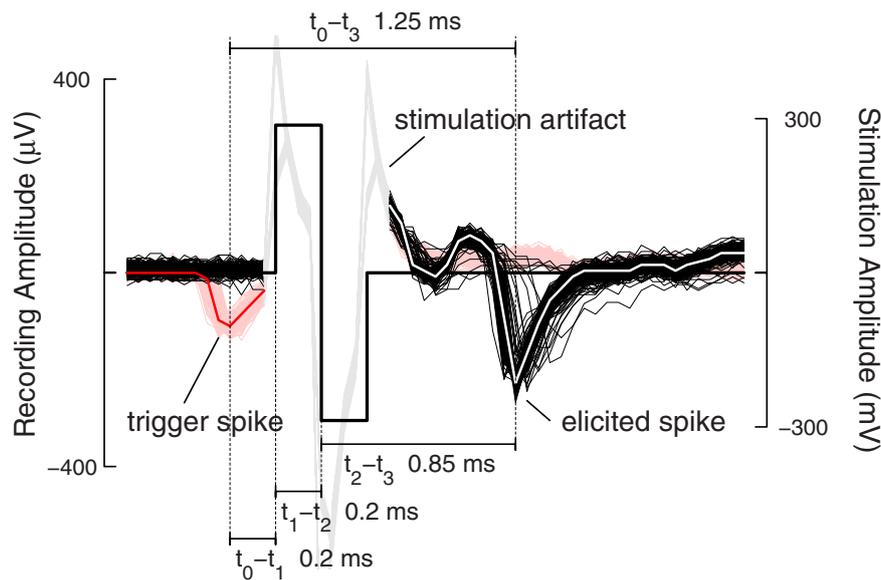


FIGURE 5 | Feedback stimulation performance. One hundred and twenty-eight traces from a closed-loop stimulation sequence are aligned at the stimulation onset-time and overlaid. Traces in red show the trigger spikes with the median over all trigger traces shown in bold red. The stimulation artifact is grayed-out for better visual clarity. The traces in black show spikes,

elicited in all but four cases after stimulation. The median over all elicited traces is shown in bold white. The antidromic propagation delay for the elicited spikes was around 0.85 ms. The different timings, detection delay, stimulation delay, and antidromic propagation delay sum up to the full loop delay of 1.25 ms.

of the stimulation pulse was $200\ \mu\text{s}$, i.e., 4 sampling periods. This delay arises as follows: $50\ \mu\text{s}$ (1 sampling period) was used to buffer the incoming data in the FPGA; $100\ \mu\text{s}$ accounted for the delay of the two-tab Butterworth filter and the last $50\ \mu\text{s}$ account for all other delays, such as synchronizing the stimulation pulse with the recording sampling time. Delays for sending digital data between the CMOS device and the FPGA were on the order of nanoseconds and thus are negligible. When stimulating with biphasic voltage pulses, the steep negative transition, which injects negative current ($I = C \times dV/dt$), is the time point, when a cell is activated (Wagenaar et al., 2004; Bakkum et al., 2008b). Thus, this time point was taken to measure the latency between stimulation and an elicited spike. In the case depicted in **Figure 5**, this timing is 0.85 ms, and the overall latency between trigger neuron activity and a spike on the elicited neuron was 1.25 ms.

As can be seen in **Figure 5**, besides achieving short feedback cycles, another advantage of using digital hardware (in this case FPGAs) for feedback generation is that no additional jitter is introduced, as such a system is fully deterministic. Sources of jitter in other systems (Hafizovic et al., 2007; Rolston et al., 2010) that close the feedback-loop around general-purpose or real-time personal computers are, for example, system interrupts that might disrupt the data processing, or buffer sizes of the USB, TCP/IP, or DAQ cards, which have to be set large enough in order to guarantee full data throughput. Usually these buffers have a size larger than one sample period. Depending on when an event happened inside this buffer, the latency could be larger or smaller and thus introduce jitter. This can be avoided by using digital hardware to hijack the data stream. In our case, the jitter was below

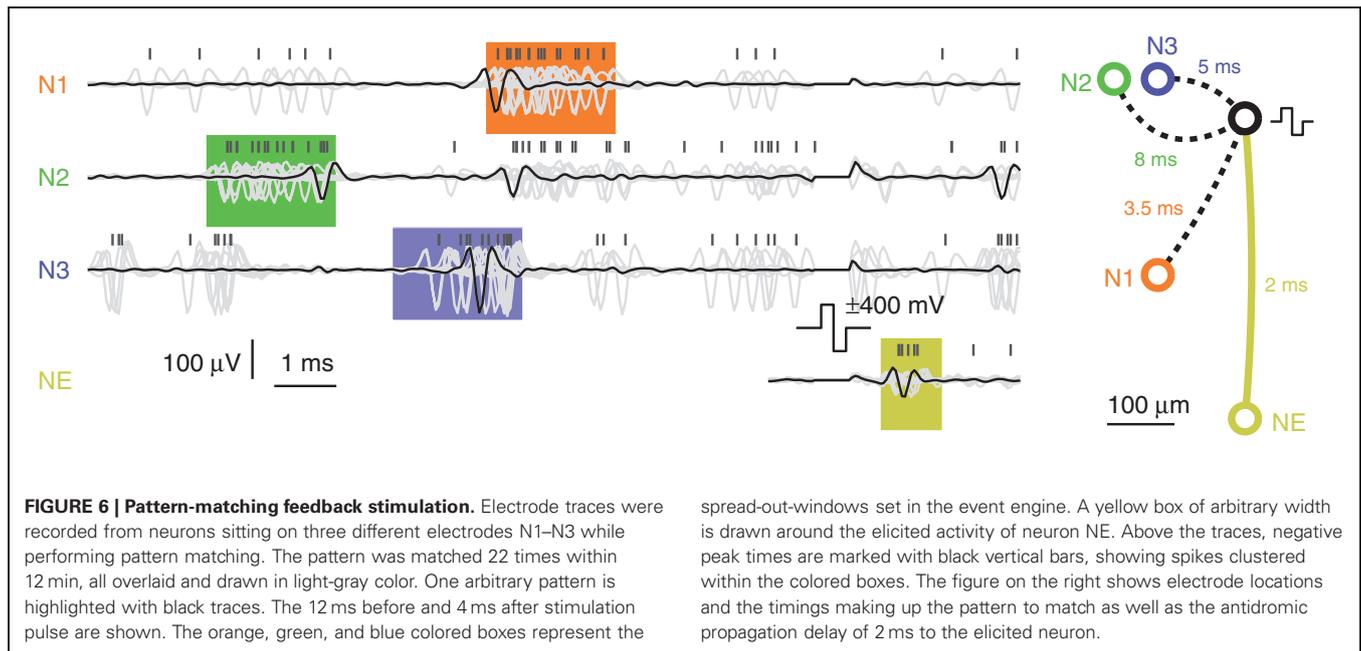
$\pm 50\ \mu\text{s}$ and arose from the fact that neural activity is, of course, not aligned to the sampling period of the CMOS-MEA ($50\ \mu\text{s}$). The exact time of the threshold crossing relative to the negative spike peak depends, among other things, on the slope of the spike waveform. Since the recorded signal was not interpolated between samples, this was an unavoidable source for jitter.

PATTERN MATCHING

To demonstrate the event engine in operation, feedback stimulation, triggered by an activity pattern, was performed. For the dataset presented in **Figure 6**, the event engine was programmed according to **Figure 3H** and classified spontaneous activity patterns as follows: A neuron recorded on electrode N2 fires an action potential; then an action potential is recorded from a neuron on electrode N3 after 3 ms; finally an action potential is recorded on electrode N1 after another 1.5 ms. Each individual event occurrence was allowed to have a jitter of $\pm 1\ \text{ms}$. After successful identification of such a pattern, a stimulation pulse was emitted to elicit action potentials on a different neuron, NE. The cell cultures under investigations typically expressed bursting behavior, and this was when almost all of the patterns occurred. During bursts, the cells usually fired more than once at an elevated frequency, and this explains why the neurons on electrodes N1–N3 showed additional spikes “outside” of the detected pattern. Nevertheless, the pattern matching event engine identified 22 activity pattern occurrences during 12 min of recording.

CORRELATION ANALYSIS

To assess the connectivity between different neurons and the efficacy of change, induced by the closed-loop feedback

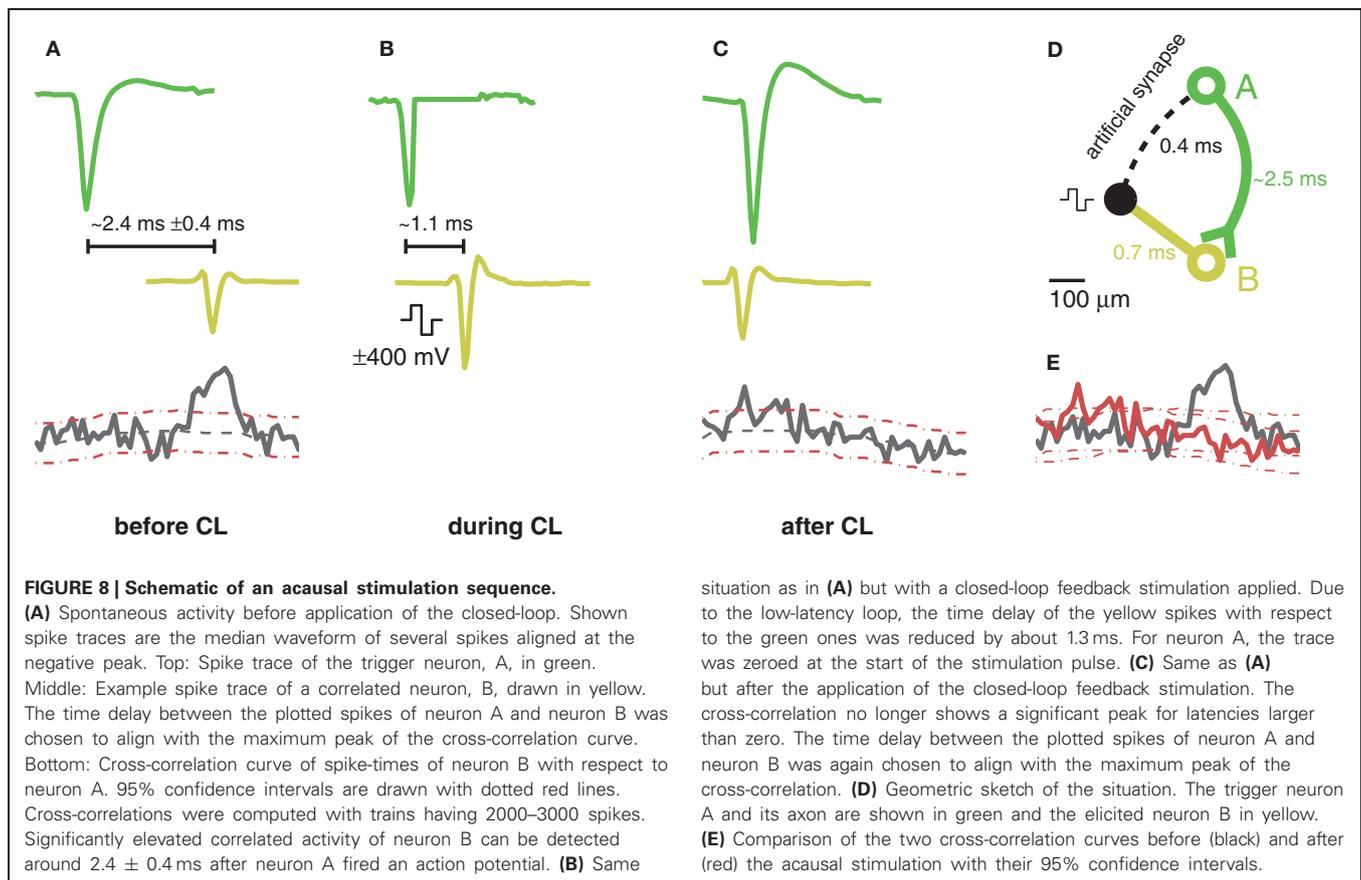
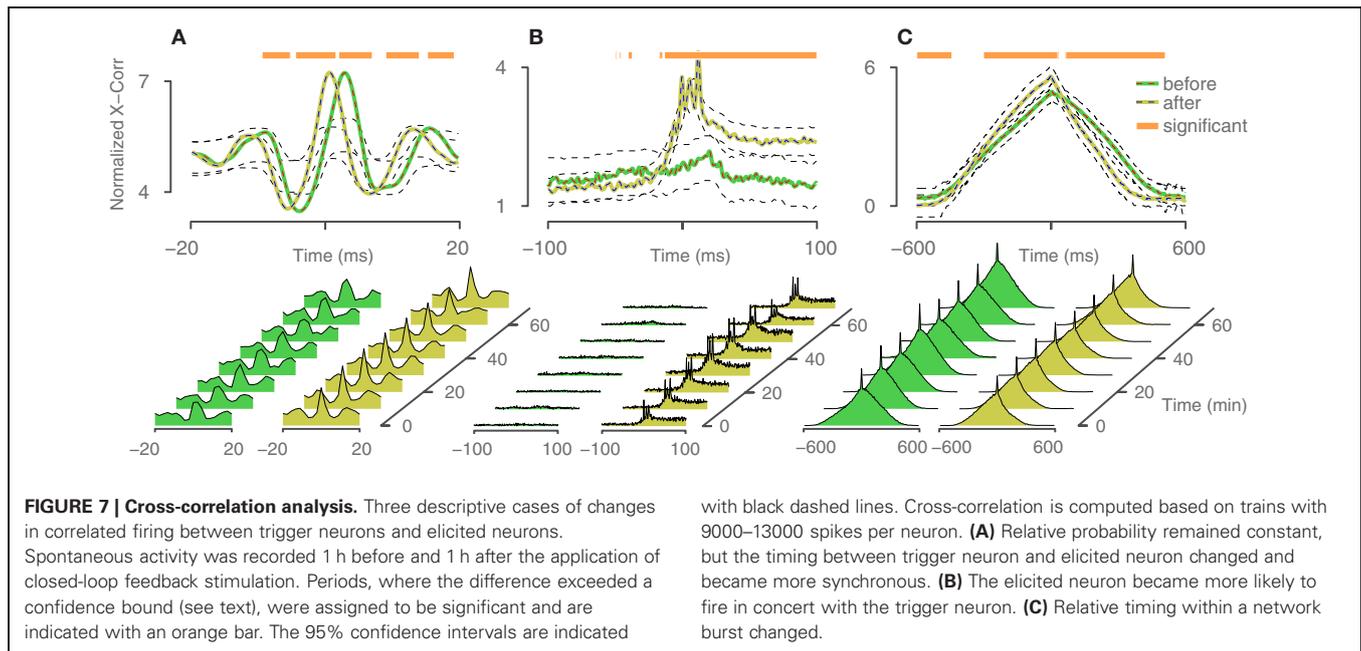


stimulation, cross-correlation curves (Perkel et al., 1967) were computed between spike trains of the trigger neuron and the elicited neurons. When exceeding a 95% confidence interval (Brillinger, 1976), correlation is considered significant. **Figure 7** shows three descriptive cases, comparing the cross-correlation curves from 1 h of spontaneous activity before and after closed-loop feedback stimulation was applied for 1 h. To evaluate significance of the change, a similar procedure as in Fujisawa et al. (2008) was used. Briefly, the two times 1 h of spontaneous activity recordings were divided into smaller bins of 10 min duration and were randomly assigned to be before or after the closed-loop stimulation. Cross-correlation from this shuffled data was computed for both “before” and “after” and the difference was evaluated. This procedure was repeated 1000 times to generate a surrogate data set. Points on the x-axis, where the true difference is larger than 95% of the surrogate data, were assigned to be significant and are marked with an orange bar in **Figure 7**. Assessing the true connectivity of neuronal networks by means of extracellular measurements is difficult, and using the cross-correlation to that end is not ideal, as effects like common inputs or firing rate changes cannot be easily explained. However, in our context of evaluating the effect of feedback stimulation, we do not necessarily seek to precisely explain the changes in network connectivity, but to rather demonstrate that a change occurred at all and to what extent.

ACAUSAL STIMULATION

One motivation for very short feedback cycles is to open the possibility of acausal stimulation. If the closed-loop stimulation ($t_0 - t_2$) is faster than the time it takes the action potential to travel along the axon and hit the synapses ($t_0 - t_3$), acausal stimulation and, therefore, induction of LTD by means of closed-loop feedback stimulation is possible. The time that it takes for an action potential, initiated at the axonal hillock, to propagate

down the axonal arbor to the synapses depends on the propagation velocity of action potentials along axons and the length of the axons. Action potential conduction velocities in unmyelinated axons were reported around $0.2\text{--}0.4\text{ ms}^{-1}$ (Debanne et al., 2011). As demonstrated in **Figure 5**, the closed-loop stimulation ($t_0 - t_2$) can be as fast as 0.4 ms, meaning acausal stimulation is possible for trigger neurons (t_0) with unmyelinated axons that synapse to an elicited neuron (t_3/s) after a minimum axial length of 80–160 μm . **Figure 8** shows such an acausal stimulation procedure. First, before applying a closed-loop, the activity between different neurons was measured then evaluated by computing the cross-correlation. In the example in **Figure 8**, the firing activity of the second neuron B with respect to the first neuron A was elevated around a delay of 2.5 ms, implying neuron A has a functional connection with neuron B. Integrating the cross-correlation curve, where it exceeds the confidence intervals around 2–3 ms after the reference time zero, reveals an integral probability of around 40% chance for neuron B to spike 2–3 ms after neuron A had fired. Once two such neurons could be identified, closed-loop stimulation can be applied between them with a very short feedback cycle. In the presented example, the delay from the trigger neuron to the elicited spike was around 1 ms, smaller than the average delay between the occurrence of their spontaneous action potentials. The closed-loop feedback stimulation was applied for 20 min, and, afterwards, the correlation was measured again. Now, the correlation no longer exceeded the confidence intervals at around 2–3 ms after the trigger neuron. Note, however, that Bi and Poo (1998) have shown that LTD can only be induced, if the spontaneous synaptic efficiency is not strong enough to evoke a post-synaptic action potential. Otherwise, the post-synaptic Ca^{2+} influx dominates, and LTP will occur. For the experiment shown in **Figure 8**, the elicited neuron spiked only a fraction of the time, and provided an intermediary synapse; in all



other cases, evoked excitatory post-synaptic currents (EPSCs) remained below the threshold. Further experiments are required before drawing conclusions. Additionally, to explore LTD and LTP in more depth, and advantageously, across many synapses

simultaneously, extracellular recordings targeted to many trigger neurons, and an elicited neuron on the CMOS-MEA could be combined with an intracellular patch-clamp, attached to the elicited neuron and measuring the incoming EPSCs.

DISCUSSION

With the presented system, capable of applying multiple flexible feedback-loops simultaneously, many different experiments will be possible. The dynamic clamp technique proved to be a valuable tool for investigating the membrane dynamics involved in action potential generation (Destexhe and Bal, 2009; Economo et al., 2010). In such systems, intracellularly applied closed-loop-controlled voltage feedback enables the manipulation of cell membrane functions. Similarly, extracellularly applied closed-loop stimulation feedback, as presented in this work, might provide a useful tool for investigating cellular and network level plasticity and enable the manipulation of neuronal network functions. Potential questions include how information processing and the amount of memory that can be stored in a cultured network are influenced by adding one or more feedback-loops. Further experiments might involve more detailed studies of both LTP and LTD of individual sets of neurons by implementing causal and acausal feedback-loops between them. Using the pattern matching capabilities of the event engine will allow for extending plasticity studies to the network level. For example, investigations of the temporal order and history of spike trains, similar to those reported by Froemke and Dan (2002) and Ikegaya et al. (2004), could be performed, however, in parallel on multiple different neurons and pathways and, in addition, the respective pathways could be dynamically altered by targeted closed-loop feedback stimulations. Further rules governing plasticity beyond the classical STDP could be investigated.

An inherent limitation of extracellular recording systems is the inability to directly measure EPSCs. Conventional plasticity studies rely on patch-clamp to directly measure the EPSC to assess synaptic connectivity strength. Since these currents are not accessible with extracellular measurement techniques, indirect methods to assess synaptic connectivity have to be employed. Although cross-correlation seems attractive and is commonly used to assess connectivity, either between different brain regions or networks, or even between individual cells, it remains to be investigated to what extent correlation analysis unveils the direct synaptic strength between neurons. A combination of patch-clamp techniques and MEAs would provide a more direct way to measure the EPSC than through the computation of cross-correlation curves. By patching the post-synaptic neuron, EPSC strengths can be directly measured and related to extracellularly recorded pre-synaptic activity. Combining the advantages of both techniques, i.e., the precise EPSC measurements through patch-clamp, and the large-scale parallel, extracellular measurements and stimulations through CMOS-MEAs with flexible feedback-loops programmed by the event engine, would greatly expand experimental horizons. One could study the plasticity

of hundreds of synapses in parallel. Furthermore, by hooking up the patch-clamp system to the event engine through dedicated spike-detection and stimulation modules, feedback-loops could be applied through the patch-clamp between extracellularly recorded and intracellularly stimulated (or vice versa) neurons.

Although, due to the high-density of electrodes, potentially all neurons can be read out individually, the recorded signals from two different neurons, located close to each other, are sometimes difficult to separate. A spike-sorting step, incorporated prior to event detection, can help to sort, and separate even neurons recorded from with the same electrodes. This holds in particular for using high-density electrode arrays (Franke et al., 2012). The spike-sorting might enable the identification of neurons with smaller spiking amplitudes, close to the noise level, and the identification of more neurons or cell assemblies. However, a drawback of more sophisticated spike-sorting algorithms is an additional time delay in the detection phase ($t_0 - t_1$). Spike-sorting, together with intracellular stimulation through patch-clamp as described above, could eliminate the aforementioned limitations in section “Recording/stimulation selectivity”: Trigger spikes can be assigned to an individual neuron through spike-sorting, and stimulation pulses will only activate action potentials in the patched neuron.

CONCLUSION

By using an FPGA to perform signal-processing, as well as feedback generation, fast, and flexible loop cycles have been realized. Our approach using reconfigurable digital hardware to perform computationally intensive tasks, such as signal filtering, spike identification, decision-making, and feedback generation, is a compromise between traditionally employed methods either using a general-purpose (micro-) processor, which introduces additional latencies, and jitter, and the highly integrated application-specific circuits (VLSI ASICs), which are much less flexible in terms of adaptations to new experimental paradigms. Our achieved closed-loop feedback latencies are lower than many axonal propagation delays and thus enable acausal stimulation. Due to the flexible event engine, high-throughput experiments applying many feedback-loops in parallel are conceivable.

ACKNOWLEDGMENTS

We thank Milos Radivojevic and Marta Lewandowska for culturing assistance and Felix Franke, Michele Fiscella, Ian Jones, and David Jäckel for helpful discussions. This work was financially supported through the ERC Advanced Grant 267351 “NeuroCMOS” and the Swiss National Science Foundation Ambizione Grant PZ00P3_132245.

REFERENCES

- Abeles, M., and Gerstein, G. L. (1988). Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *J. Neurophysiol.* 60, 909–924.
- Azin, M., Guggenmos, D. J., Barbay, S., Nudo, R. J., and Mohseni, P. (2011). A battery-powered activity-dependent intracortical microstimulation ic for brain-machine-brain interface. *IEEE J. Solid-State Circ.* 46, 731–745.
- Bakkum, D. J., Chao, Z. C., and Potter, S. M. (2008a). Long-term activity-dependent plasticity of action potential propagation delay and amplitude in cortical networks. *PLoS ONE* 3:e2088. doi: 10.1371/journal.pone.0002088
- Bakkum, D. J., Chao, Z. C., and Potter, S. M. (2008b). Spatio-temporal electrical stimuli shape behavior of an embodied cortical network in a goal-directed learning task. *J. Neural Eng.* 5, 310–323.
- Berdondini, L., Imfeld, K., Maccione, A., Tedesco, M., Neukom, S., Koudelka-Hep, M., et al. (2009). Active pixel sensor array for high spatio-temporal resolution electrophysiological recordings

- from single cell to large scale neuronal networks. *Lab Chip* 9, 2644–2651.
- Bi, G., and Poo, M. (1999). Distributed synaptic modification in neural networks induced by patterned stimulation. *Nature* 401, 792–796.
- Bi, G. Q., and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.
- Bienenstock, E. (1995). A model of neocortex. *Netw. Comput. Neural Syst.* 6, 179–224.
- Bontorin, G., Renaud, S., Garenne, A., Alvado, L., Le Masson, G., and Tomas, J. (2007). A real-time closed-loop setup for hybrid neural networks. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2007, 3004–3007.
- Brillinger, D. R. (1976). Measuring the association of point processes: a case history. *Am. Math. Mon.* 83, 16–22.
- Chang, E. Y., Morris, K. F., Shannon, R., and Lindsey, B. G. (2000). Repeated sequences of interspike intervals in baroresponsive respiratory related neuronal assemblies of the cat brain stem. *J. Neurophysiol.* 84, 1136–1148.
- Chen, T.-C., Chen, K., Yang, Z., Cockerham, K., and Liu, W. (2009). “A biomedical multiprocessor SoC for closed-loop neuroprosthetic applications,” in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*. (San Francisco, CA).
- Debanne, D., Campanac, E., Bialowas, A., Carlier, E., and Alcaraz, G. (2011). Axon physiology. *Physiol. Rev.* 91, 555–602.
- Destexhe, A., and Bal, T. (eds.). (2009). *Dynamic-Clamp: From Principles to Applications*. Springer Series in Computational Neuroscience. Vol. 1. XIV, 429, 60 illus. ISBN: 978-0-387-89278-8.
- Economo, M. N., Fernandez, F. R., and White, J. A. (2010). Dynamic clamp: alteration of response properties and creation of virtual realities in neurophysiology. *J. Neurosci.* 30, 2407–2413.
- Franke, F., Jäckel, D., Dragas, J., Müller, J., Radivojevic, M., Bakkum, D. J., et al. (2012). High density microelectrode array recordings and real-time spike sorting for closed-loop experiments: an emerging tool to study neural plasticity. *Front. Neural Circuits* 6:105. doi: 10.3389/fncir.2012.00105
- Frey, U., Sedivy, J., Heer, F., Pedron, R., Ballini, M., Mueller, J., et al. (2010). Switch-matrix-based high-density microelectrode array in CMOS technology. *IEEE J. Solid-State Circ.* 45, 467–482.
- Froemke, R. C., and Dan, Y. (2002). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature* 416, 433–438.
- Fujisawa, S., Amarasingham, A., Harrison, M. T., and Buzsáki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat. Neurosci.* 11, 823–833.
- Hafizovic, S., Heer, F., Ugnienko, T., Frey, U., Blau, A., Ziegler, C., et al. (2007). A CMOS-based microelectrode array for interaction with neuronal cultures. *J. Neurosci. Methods* 164, 93–106.
- Hales, C. M., Rolston, J. D., and Potter, S. M. (2010). How to culture, record and stimulate neuronal networks on micro-electrode arrays (MEAs). *J. Vis. Exp.* 39:e2056. doi:10.3791/2056
- Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., et al. (2004). Synfire chains and cortical songs: temporal modules of cortical activity. *Science* 304, 559–564.
- Izhikevich, E. M. (2006). Polychronization: computation with spikes. *Neural Comput.* 18, 245–282.
- Izhikevich, E. M., Gally, J. A., and Edelman, G. M. (2004). Spike-timing dynamics of neuronal groups. *Cereb. Cortex* 14, 933–944.
- Jackson, A., Mavoori, J., and Fetz, E. E. (2006a). Long-term motor cortex plasticity induced by an electronic neural implant. *Nature* 444, 56–60.
- Jackson, A., Moritz, C. T., Mavoori, J., Lucas, T. H., and Fetz, E. E. (2006b). The Neurochip BCI: towards a neural prosthesis for upper limb function. *IEEE Trans. Neural Syst. Rehabil. Eng.* 14, 187–190.
- Lee, J., Rhew, H.-G., Kipke, D. R., and Flynn, M. P. (2010). A 64 channel programmable closed-loop neurostimulator with 8 channel neural amplifier and logarithmic, ADC. *IEEE J. Solid-State Circ.* 45, 1935–1945.
- Lipski, J. (1981). Antidromic activation of neurones as an analytic tool in the study of the central nervous system. *J. Neurosci. Methods* 4, 1–32.
- Livi, P., Heer, F., Frey, U., Bakkum, D. J., and Hierlemann, A. (2010). Compact voltage and current stimulation buffer for high-density microelectrode arrays. *IEEE Trans. Biomed. Circuits Syst.* 4, 372–378.
- Markram, H., Lubke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275, 213–215.
- Mavoori, J., Jackson, A., Diorio, C., and Fetz, E. (2005). An autonomous implantable computer for neural recording and stimulation in unrestrained primates. *J. Neurosci. Methods* 148, 71–77.
- Novellino, A., D’Angelo, P., Cozzi, L., Chiappalone, M., Sanguineti, V., and Martinoia, S. (2007). Connecting neurons to a mobile robot: an *in vitro* bidirectional neural interface. *Comput. Intell. Neurosci.* 2007, 13.
- Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967). Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys. J.* 7, 419–440.
- Rebesco, J. M., Stevenson, I. H., Kording, K. P., Solla, S. A., and Miller, L. E. (2010). Rewiring neural interactions by micro-stimulation. *Front. Syst. Neurosci.* 4:39. doi: 10.3389/fnsys.2010.00039
- Rizk, M., Bossetti, C. A., Jochum, T. A., Callender, S. H., Nicoletis, M. A., Turner, D. A., et al. (2009). A fully implantable 96-channel neural data acquisition system. *J. Neural Eng.* 6:026002. doi: 10.1088/1741-2560/6/2/026002
- Rolston, J. D., Gross, R. E., and Potter, S. M. (2010). Closed-loop, open-source electrophysiology. *Front. Neurosci.* 4:31. doi: 10.3389/fnins.2010.00031
- Rolston, J. D., Wagenaar, D. A., and Potter, S. M. (2007). Precisely timed spatiotemporal patterns of neural activity in dissociated cortical cultures. *Neuroscience* 148, 294–303.
- Song, S., Miller, K. D., and Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci.* 3, 919–926.
- Venkatraman, S., Elkabany, K., Long, J. D., Yao, Y., and Carmena, J. M. (2009). A system for neural recording and closed-loop intracortical microstimulation in awake rodents. *IEEE Trans. Biomed. Eng.* 56, 15–22.
- Wagenaar, D. A., Pine, J., and Potter, S. M. (2004). Effective parameters for stimulation of dissociated cultures using multi-electrode arrays. *J. Neurosci. Methods* 138, 27–37.
- Wallach, A., Eytan, D., Gal, A., Zrenner, C., and Marom, S. (2011). Neuronal response clamp. *Front. Neuroeng.* 4:3. doi: 10.3389/fneng.2011.00003
- Zanos, S., Richardson, A. G., Shupe, L., Miles, F. P., and Fetz, E. E. (2011). The Neurochip-2: an autonomous head-fixed computer for recording and stimulating in freely behaving monkeys. *IEEE Trans. Neural Syst. Rehabil. Eng.* 19, 427–435.
- Zrenner, C., Eytan, D., Wallach, A., Thier, P., and Marom, S. (2010). A generic framework for real-time multi-channel neuronal signal analysis, telemetry control, and sub-millisecond latency feedback generation. *Front. Neurosci.* 4:173. doi: 10.3389/fnins.2010.00173

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 September 2012; paper pending published: 30 October 2012; accepted: 22 December 2012; published online: 10 January 2013.

Citation: Müller J, Bakkum DJ and Hierlemann A (2013) Sub-millisecond closed-loop feedback stimulation between arbitrary sets of individual neurons. *Front. Neural Circuits* 6:121. doi: 10.3389/fncir.2012.00121

Copyright © 2013 Müller, Bakkum and Hierlemann. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.