**1. Introduction:**

*What beauty is shown in the preparations obtained by the precipitation of silver dichromate deposited exclusively onto the nervous elements! But, on the other hand, what dense forests are revealed, in which it is difficult to discover the terminal endings of its intricate branching… Given that the adult jungle is impenetrable and indefinable, why not study the young forest, as we would say in its nursery stage.*

 *Santiago Ramón y Cajal (1852-1934)*

It is commonly accepted that the information processing in the brain is carried out by large groups of interconnected neurons. Neurons are the cells responsible for encoding, transmitting, and integrating signals originating inside or outside the nervous system. The transmission of information within and between neurons involves changes in the resting membrane potential, when compared to the extracellular space. The inputs one neuron receives at the synapses from other neurons cause transient changes in its resting membrane potential, called postsynaptic potentials. These changes in potential are mediated by the flux of ions between the intracellular and extracellular space. The flux of ions is made possible through ion channels present in the membrane. The ion channels open or close depending on the membrane potential and on substances released by the neurons, namely neurotransmitters, which bind to receptors on the cell’s membrane and hyperpolarize or depolarize the cell. When the postsynaptic potential reaches a threshold, the neuron produces an impulse. The impulses or spikes, called action potentials, are characterized by a certain amplitude and duration and are the units of information transmission at the interneuronal level [1]. The discovery of the neuron was a milestone in brain research and paved the way for modern neuroscience, but the brain is yet to yield the vast majority of its secrets.

The current neuroscience research operates at two disconnected levels: The macro- and microscopic levels. The macroscopic level uses imaging techniques like functional Magnetic Resonance Imaging (fMRI) and Magnetoencephalography (MEG) to measure regional changes in metabolism and blood flow associated with changes in brain activity. It captures whole brain activity patterns that allow the mapping of brain regions associated with a particular behavior or task. These techniques lack single-cell details and the requisite temporal resolution to permit detection of neuronal firing patterns. The microscopic level is concerned with investigating how individual nerve cells work, studying their response to stimulation and monitoring the firing rates associated with a certain behavioral output, mental state or motor activity. This can be done using implanted electrodes to record the rates and timing of action potentials. The sparse sampling of neuronal activity monitoring tens to few hundreds of neurons does not give the global view of signaling in neural circuits that can involve millions of neurons.

There is a gap between the two levels, that is believed to entail an answer to the question of how neuron cells collaborate to process information. To fill in the gap, we need the static anatomical map of the brain circuitry describing the synaptic connections within any given brain area, as well as the dynamic map revealing the patterns and sequences of neuronal firing by all neurons over time scales on which behavioral outputs or mental states occur. Hence the aspiration is not only to map the "impenetrable jungle" that Cajal referred to but also to map the dynamical traffic within the jungle and analyze it. Research efforts are conducted to approach that ultimate goal, and along the hard path to achieve it, technological breakthroughs evolved and more are bound to arise. New technologies may include new optical techniques to image in 3D, new capabilities for storage and manipulation of massive data sets, new clinically viable Brain Machine Interfaces to help paralyzed patients and development of biologically inspired computational devices.[2]

Focusing on the Microscopic level, the following are two of the research fields concerned with recordings of the spiking activity of neurons using microelectrode arrays.

**(a) Brain-Machine Interface:**

Extracting motor control signals from the firing patterns of populations of neurons and using these control signals to reproduce motor behaviors in artificial actuators are the key operations of Brain-Machine Interface (BMI) [3]. The typical neural signal processing pathway as shown in fig.1 is designed to measure the instantaneous frequency of neural action potentials, or spikes. Since any given electrode may sense spikes from multiple neurons, it is typically necessary to sort all detected spikes by wave shape (i.e. by neuron). Firing rates of sorted spikes are typically measured by moving average; these rates can then be used by “decoding” algorithms which use statistical models to correlate spiking activity with behavioral or motor activity in the subject.

**Fig.1.1**  Block diagram of the typical pathway of brain machine interface

Hence invasive BMIs rely on the physiological property of individual cortical neurons to modulate their spiking activity in association with movements [3]. These modulations are found to be highly variable from neuron to neuron and from trial to trial. Yet averaging across many trials reveals fairly consistent firing patterns. Based on the hypothesis that the function of neural circuits is an emergent property that arises from the coordinated activity of large numbers of neurons, this phenomenon can be explained. Individual neurons generally form synaptic connections with thousands of other neurons. In distributed circuits, the larger the connectivity matrix the greater the redundancy within the network. Given their distributed connections and their plasticity, neurons are likely to be subject to continuous dynamic rearrangement, participating at different times in different active ensembles [2]. Accordingly both accuracy and reliability of predictions of motor activity improve considerably with increasing the number of simultaneously recorded neurons and decreasing the errors due to individual neuron firing variability. Pursuing this motivation, the number of simultaneously recorded neurons has been approximately doubling every 7 years since 1950’s [4]. Standard recording techniques using 704 implantable micro wire arrays have been reported in literature [5]. Recently Nicolelis lab at Duke University announced their success to simultaneously record the electrical activity produced by a population of 1,874 interconnected single neurons at work in a primate.

**(b) Brain in a Dish:**

At present, the prime methodology for studying neuronal extracellular activity under in vitro conditions is by using substrate-integrated microelectrode arrays (MEAs). This methodology permits simultaneous, long-term recordings (i.e. of up to several weeks) of extracellular field potentials. Correlating MEA recordings with microscopic imaging and stimulations is widely used to study the circuit-connectivity, dynamics and propagation effects in neuron assemblies. It is also used to investigate population coding, activity patterns, plasticity and pharmacological testing on either dissociated neuronal cultures or brain slices of embryonic rats, i.e. the young forests as Cajal described them.

Commercially available MEA systems integrate typically 60–120 microelectrodes of 10–30 *μ*m in diameter with pitches on the order of hundreds of micrometers. Typical neuron soma dimension in vertebrates is few micrometers long and the typical neuronal networks have 10000–50000 neurons, the limited number of electrodes and their rather large pitch results in a substantial spatial undersampling of the overall network activity [6] as shown in fig2.

**Fig. 2:** Substrate-integrated MEA dish. The microscopic image of the electrode (black) and neurons

The development of higher spatial and temporal resolution at low noise levels are prerequisites for opening the perspective to access the network electrical activity at the global and cell levels. Recently, CMOS-based high-density MEAs were developed featuring switching techniques to manage a large number of electrode channels interconnections, multiplexing, amplification, and filtering. Active Pixel Sensor based MEA platform providing 4096 microelectrodes at 21µm inter-electrode separation and 7.7KHz sampling rate has been documented [6].

Considering the ultimate goal of Brain Activity Map [2], the current neuroscience in vivo and in vitro research states and the advancement of high density microelectrode arrays, the migration to monitoring thousands of recording channels at high temporal resolution is achievable.

**1.1 Increasing the number of Recording Channels:**

*More is Different - The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead at each level of complexity entirely new properties appear.*

  *Philip Warren Anderson*

The augmentation of the number of recording channels carries different challenges to the neural signal processing system. The primary challenge is the massive increase in recorded data that needs proactive strategies for data transfer, reduction, management and analysis. The implementation of real-time signal processing becomes essential to alleviate huge data storage requirements. The access to a more detailed view of neuronal networks might reveal new properties and challenges pushing for the development of new analyzing tools.

With the continuous advancement of data acquisition systems featuring high-count recording channels, there exists a clear need for a test bed to develop and investigate a more suitable new generation of Neural Signal Processing (NSP) algorithms and computational tools. The platform has to offer programmable flexibility to allow the trial of different new strategies and novel computational techniques as well as rigorous testing for evaluation.

 A plausible NSP platform that can handle thousands of recording channels has to provide means of high data transfer. As a numerical example, a NSP platform handling 2560 channels sampled at 31.25 KH at a sample precision of 16-bits must be capable of managing an input data stream of 1.28Gbps. The data transfer interface has to be compatible with high-density neural data acquisition systems [7].

Data reduction based on the sparse nature of the neural signal with respect to time and the redundancy perceived across multiple electrode recordings becomes essential. Spike detection is the essential first step building block that allows the system to deliver only the action potential waveforms, their respective occurrence times and channel ID instead of the entire raw signal. The AP waveforms are then used by an autonomous spike sorter to first distinguish true spikes from false detections, then, to associate each spike to its generating neuron in case of multi-unit recordings. Depending on the performance and inter-electrode spacing, the AP waveforms might be necessary to identify redundancy over multiple recording channels.

The spike detection settings for each channel is independent from the settings of other channels, and hence spike detection over different sites can run in parallel. Applying parallel processing whenever possible limits the overall latency and assists in achieving real time implementation.

The NSP platform has to be fully autonomous and functional under expected Signal-to-Noise Ratios delivered by the data acquisition system. The system must be adaptive to varying noise levels over different channels and over time.

The main objective of the proposed project is to design an experimental test bed that can facilitate dealing with a large number of recorded neurons in real time. It also presents an architecture that performs spike-based data reduction.

 **1.2. Why consider FPGA?**

Ross Freeman (1944-1989) established the leading FPGA developer Xilinx in 1984 and invented a year later the first Field Programmable Gate Array (FPGA). **FPGAs** are programmable semiconductor devices that are based around a matrix of Configurable Logic Blocks (CLBs) connected through programmable interconnects. FPGAs can be configured to implement custom hardware applications and functionalities. Since their invention, FPGAs have evolved far beyond the basic capabilities present in their predecessors, and incorporate hard Application Specific Integrated Blocks of commonly used functionality such as RAM, clock management, and DSP.

FPGAs are parallel in nature, so different processing operations do not have to compete for the same resources. Each independent processing task is assigned to a dedicated section of the chip and can function autonomously without any influence from other logic blocks.

As integrated circuits grew smaller and maximum toggle rates increased the need for input/output bandwidth exploded. With more hardware resources and faster clock speeds, conventional I/O resources became the bottleneck to FPGA performance. In 2002, Xilinx embedded high-speed serial Multi-gigabit transceivers (MGTs) on their FPGAs and introduced them commercially under the name Rocket I/O. MGTs are Serializers/Deserilizers (SERDES) that allow serial data transmission over differential pairs at speeds of up to 28.05Gbps per lane (see Fig. 1.3). Alternatively, multiple MGTs can be bonded together to form a higher bandwidth interface. Multiple MGTs are integrated above and below the Block RAM columns providing close availability for ingress and egress FIFOs. Rocket IO serial transceivers (see Fig 1.4) are compliant with standard gigabit communication protocols.

3.125

6.25

GTP, 3.125

**Fig. 1.3**. Bar graph presenting the available serial Multi-Gigabit Transceiver line rates.

FPGAs offer massive parallel processing performance, reconfigurable flexibility and superior capabilty of streaming data, and therefore present an appealing hardware implementation solution for a NSP testbed that can handle a large number of similarly strutured parallel channels in real time.



**Fig. 1.4:** Functional Block diagram of Xilinx® GTP MGT.