# **IBM Research Report**

## Scaling, Stability, and Synchronization in Mouse-sized (and Larger) Cortical Simulations

Rajagopal Ananthanarayanan, Dharmendra S. Modha

IBM Research Division Almaden Research Center 650 Harry Road San Jose, CA 95120-6099



Research Division Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties). Copies may be requested from IBM T. J. Watson Research Center, P. O. Box 218, Yorktown Heights, NY 10598 USA (email: reports@us.ibm.com). Some reports are available on the internet at <a href="http://domino.watson.ibm.com/library/CyberDig.nsf/home">http://domino.watson.ibm.com/library/CyberDig.nsf/home</a>.

#### Scaling, Stability, and Synchronization in Mouse-sized (and Larger) Cortical Simulations

#### Rajagopal Ananthanarayanan and Dharmendra S. Modha

Cognitive Computing Department, IBM Almaden Research Center, San Jose, CA 95120

Email: {ananthr@us.ibm.com,dmodha@us.ibm.com}

Together, the two hemispheres of the mouse cortex contain, roughly,  $16 \times 10^6$  neurons, and 8,000 synapses per neuron [1]. Modeling and simulation at this scale (and beyond) is an emerging area of scientific exploration that will likely lead to a host of new computational and neuro-scientific insights. To facilitate *near real-time* study of network dynamics, we have recently developed a massively parallel cortical simulator simulator [2] that incorporates relatively simpler phenomenological models of spiking neurons [3], spiketiming dependent plasticity [4], and axonal conductance delays. In this paper, we explore the challenges involved in scaling, stabilizing, and synchronizing such large-scale models.

#### Scaling

We created a basic connection template with 32, 768 "groups" each of which have 500 neurons. Each group sends axons to 100 randomly selected groups and each neuron from the projecting group makes a total of c = 80 synapses with the neurons of the receptive group. The resultant network consisted of a total of 16, 384,000 neurons and 8,000 synapses per neuron comparable to a mouse. Roughly, 80% of the groups were excitatory and the remaining were inhibitory. For excitatory groups we used axonal delays uniformly ranging from 1-20 ms, and for inhibitory groups we used a fixed 1ms axonal delay. This network structure is not intended to capture detailed anatomy of the mouse brain, but is a step towards understanding dynamics of large scale models.

Using a BlueGene/L supercomputer with 8, 192 processors (2 times the size in [2]), with 4 TB of main memory (4 times the size in [2]), we were able to load the entire network and all the associated auxiliary data structures. In a stable, sustained firing mode (see below), using a stimulus pattern that delivered super-threshold excitation to every neuron at 4Hz, using 1ms simulation time steps, we were able to simulate 5s of model time in 168s of real-time at a mean firing rate of 4.95Hz. Effectively, when normalized to a 1Hz mean firing rate, we were able to run a 1s of model time in approximately 7s of real-time. Interestingly, with everything else kept the same, but without using spike-timing dependent plasticity, we were able to run 5s of model time in 75s of real-time. When normalized to a 1Hz mean firing rate, this amounts to 1s of model time in approximately 3s of real-time.

To further push the boundaries of scaling, by using c = 160 above, we created a network with 16, 384, 000 neurons and 16,000 synapses per neuron (alternatively, we could have created a network with 32,768,000 neurons and 8,000 synapses per neuron). We were able to run this model on a larger BlueGene/L supercomputer with 16,384 processors and 8 TB of main memory. Specifically, in a stable, sustained firing mode, when using a 5Hz stimulation, we were able to achieve 5s of model time in 265s of real-time at a mean firing rate of 5Hz. When normalized to a 1Hz mean firing rate, this amounts to 1s of model time in approximately 11s of real-time.

#### **Stability and Synchronization**

Stabilizing cortical simulations is enormously difficult: "For a large network of excitatory and inhibitory



Figure 1: Dynamics of a network with 16, 384, 000 neurons and 16, 000 synapses per neuron. Panels (a), (b), and (c) display percetage of neurons that fired at each simulation time step. Panels (a), (b), and (c) exhibit, respectively, damped, sustained, and avalanche behaviors. Panel (d) shows the raster of 1,000 neurons (out of 16, 384, 000 neurons) corresponding to stable mode in panel (b).

neurons with small EPSPs it is very difficult, if not impossible, to attain steady ongoing activity at low firing rates." [5] It is very easy to drive a network into avalanche mode or into an uninteresting quiescent state, but is difficult to maintain it in a stable periodic, quasi-periodic, or chaotic states [6].

A neuron decides to fire on the basis of co-incident activity at its synapses and the synaptic efficacies. We found that the allowed *maximum synaptic efficacy* (which upper bounds the growth of excitatory synaptic efficacies under spike-timing dependent plasticity) and the *probability of the super-threshold stimulus* together greatly affected the behavior of our networks described above. During the course of this study, we used several models with varying number of synapses from 1 synapse per neuron to 16,000 synapses per neuron. We observed that finding a range of maximum synaptic efficacies corresponding to stable models is harder to achieve for higher number of synapses per neuron if the stimulus probability is kept low. Further, there appears to be a threshold stimulus probability below which – when maximum synaptic efficacy is varied – models make a sharp transition from damped mode to avalanche mode. For the network with 16,000 synapses per neuron, the three modes, namely, damped, sustained, and avalanche, are captured in Figure 1. We found similar behavior also for the smaller, mouse-scale network with 8,000 synapses per neuron.

### References

- [1] Braitenberg V, Schüz A: Cortex: Statistics and Geometry of Neuronal Conectivity. Springer 1998.
- [2] Frye J, Ananthanarayanan R, Modha DS: Towards real-time, mouse-scale cortical simulations. In *CoSyNe: Computational and Systems Neuroscience, Salt Lake City, Utah* 2007.
- [3] Izhikevich EM, Gally JA, Edelman GM: Spike-Timing Dynamics of Neuronal Groups. *Cerebral Cortex* 2004, **14**:933–944.
- [4] Markram H, Lubke J, Frotscher M, Sakmann B: Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 1997, 275:213–215.
- [5] Abeles M: Corticonics: Neural Circuits of the Cerebral Cortex. Cambridge University Press 1991.
- [6] Vodels TP, Rajan K, Abbott LF: Neural network dynamics. Annu. Rev. Neuroscience 2205, 28:357– 376.