2016

Graceful Degradation Under Noise on Brain Inspired Robot Controllers

de Azambuja, R

http://hdl.handle.net/10026.1/10539

Graceful Degradation under Noise on Brain Inspired Robot Controllers

Ricardo de Azambuja[1,2], Frederico B. Klein[1], Martin F. Stoelen[1], Samantha V. Adams[1], and Angelo Cangelosi[1]

[1] School of Computing, Electronics and Mathematics,
Plymouth University, Plymouth, UK
[2] CAPES Foundation,
Ministry of Education of Brazil, Brasilia, DF 70040-020, Brazil
{ricardo.deazambuja, frederico.klein, martin.stoelen,
samantha.adams, a.cangelosi}@plymouth.ac.uk
http://www.plymouth.ac.uk
http://www.capes.gov.br

Abstract How can we build robot controllers that are able to work under harsh conditions, but without experiencing catastrophic failures? As seen on the recent Fukushima's nuclear disaster, standard robots break down when exposed to high radiation environments. Here we present the results from two arrangements of Spiking Neural Networks, based on the Liquid State Machine (LSM) framework, that were able to gracefully degrade under the effects of a noisy current injected directly into each simulated neuron. These noisy currents could be seen, in a simplified way, as the consequences of exposition to non-destructive radiation. The results show that not only can the systems withstand noise, but one of the configurations, the Modular Parallel LSM, actually improved its results, in a certain range, when the noise levels were increased. Also, the robot controllers implemented in this work are suitable to run on a modern, power efficient neuromorphic hardware such as SpiNNaker.

Keywords: SNN, Liquid State Machines, robot control, noise, graceful degradation, robustness

1 Introduction

Five years have passed since Fukushima's nuclear disaster and current robot technology is still not ready for such a big challenge. The high level of radiation in areas close to the reactors was lethal for human beings and the robots sent to the site have severely suffered from it, hence making clear the need for more research. Modern computers, and therefore robot controllers, are designed around digital circuits and, despite several advances in manufacturing processes, design and simulation, they are still not immune to it. Digital systems also suffer from non-destructive radiation, since it can generate Single-Event Upsets (SEU) or "soft-errors". A SEU is an alteration in a logic state as a result of an energetic

particle entering the microelectronic device [11]. In addition to man-made radiation sources, space and terrestrial environments are also subjected to cosmic rays and naturally available radioactive isotopes.

There is evolutionary pressure for natural information processing systems to be fault tolerant. If damaged or malfunctioning neuronal cells were to change drastically the overall behaviour of the organism, it would restrict chances of survival. According to [16], graceful degradation is defined as "graded, probabilistic deficits, with some sparing of function, and with performance strongly influenced by the frequency or familiarity of the stimulus and/or its degree of consistency with other items". As such, we can identify graceful degradation in a number of neural systems. For example, in motor control neurologic disorders, Essential Tremor (characterised by periodic 4-12Hz low amplitude movements) is among the most common (prevalence $\approx 4.0\%$ among aged 40 years or older), while motor disorder with choreoathetotic and ballistic movements i.e. "excessive, spontaneous, irregularly timed, non-repetitive, randomly distributed and abrupt in character" has a prevalence $\approx 0.01\%$ [9,2,22]. As another example, in memory encoding and consolidation, engram cells distribute learned information so that no individual neuron is responsible for a particular information, but their collective activation. Recent advances are being made in elucidating the biology behind this [19,18].

Efficiency is another characteristic seen all around nature designs. The best example is the human brain, since cortex and cerebellum together spend on average around 15W [6]. On the other hand, the Human Brain Project expects to simulate the whole brain, in the cellular level, using an exascale computer or 60MW [12].

In an attempt to start developing solutions for the current problems robotic systems encounter when exposed to an environment with a high level of radiation, we propose in this work the use of biologically inspired robot controllers [1] for a more nature-like graceful degradation, instead of a catastrophic failure, when exposed to it. Modular and Monolithic designs of a special type of feedback enhanced parallel Liquid State Machines (LSM) [15,14] are exposed to different noise levels, in a simulated environment, and the results analysed with a robotic task as the benchmark. White Gaussian noise is injected directly into the neuron model, which could be seen as an example of the result from the non-destructive effects of radiation. Additionally, LSM are modelled based on Spiking Neural Networks (SNN), therefore power efficiency could be easily acchieved implementing the SNN in a neuromorphic hardware such as SpiNNaker [5], BrainScaleS [21] or Silicon Neurons (SiN) [7] which could also improve the reliability even further.

2 Materials and Methods

The investigation presented here was based on an earlier work[3], where a new humanoid robot control framework using parallel, diverse and noisy groups of

---
[3] Source code available at github.com/ricardodeazambuja/IJCNN2016

biologically inspired LSM was introduced [1]. This robot controller was able to reproduce trajectories (shapes) previously learned from a teacher, but the effects of varying noise levels were not studied.

In this new work, eleven different noise levels (100 trials each), starting from the standard one defined in [8] and going up to 100% above that (see Section 2.2), were employed to verify the noise effect on two different parallel LSM configurations: Modular and Monolithic (see Section 2.1). The final analysis was done through the robot's resultant movement performing the benchmark task of drawing a square shape on a table (see Section 2.3).

2.1   Modular and Monolithic Parallel LSM

The idea of breaking an LSM into multiple liquids (or simplified models of cortical columns) in parallel to increase the computational power was initialy presented in [15], but only in [1] was an external feedback loop, as suggested in [8,14], explored for this particular situation. Also the parallel system presented in [15] had an external output layer (*readout*) shared among all neurons contrasting with the one presented in [1] where each liquid was trained individually and had its own *readout* resulting in a system with improved learning capabilities. Those two approaches are called here the *Monolithic Parallel LSM* (Figure 1b) and *Modular Parallel LSM* (Figure 1a), respectively. To facilitate comparisons, the same random seeds from [1], therefore the same *liquids*, were employed here, but the readout layers were trained again as the Monolithic approach has not been tested before.
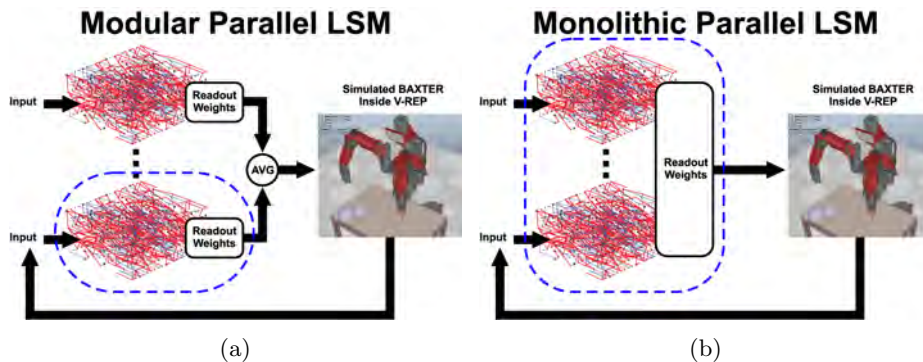

Figure 1: The Modular approach (a) uses individual readout layers for each liquid. The Monolithic approach (b) has only one readout layer shared among all its neurons. Both systems reuse the same five LSM (liquids) from [1], but with retrained readouts.

2.2 Neuron Model and Noise Levels

The neuron model applied in this work, the Leaky Integrate and Fire (LIF) partially represented by the Equation 1 (for more details see [1]), has its membrane reset voltage (V_{reset}) drawn from a uniform distribution ($[13.8mV, 14.49mV]$) when the neural network is created and generates a spike when it reaches $15mV$ ($V_{threshold}$). On the algorithmic level, the membrane voltage is always clamped between $-15mV$ and $+15mV$, although its rest potential is $0mV$ and it is set back to the reset value (V_{reset}) after every spike. Consequently, most of the time, the neuron membrane will fluctuate between V_{reset} and $V_{threshold}$ or, in the worst scenario, with $\Delta V \approx 1.2mV$.

The simulation of a faulty system through the injection of noise (see Section 1) is accomplished using the i_{noise} variable from Equation 1. Its value is drawn from a Gaussian distribution ($\mu = 0$ and $\sigma = 1nA$) multiplied accordingly to what we call here *noise level* (A_{noise}). Having a noise level of 100%, 110%, 120%, ..., 200% means the multiplier value goes from 1.0 up to 2.0. The parameters were defined according to what was presented in [15] and [8], hence $c_m = 30nF$ and $\tau_m = 30ms$. This yields, ignoring other noise sources, a Signal-to-noise ratio (SNR) of approximately $\left(\frac{\Delta V/mV}{A_{noise}}\right)^2$. Thus the system has its SNR varied from 1.44 to 0.36 (see Figure 2).

$$\frac{dv(t)}{dt} = \frac{i_e(t) + i_i(t) + i_{offset} + i_{noise}(t)}{c_m} + \frac{v_{rest} - v(t)}{\tau_m} \tag{1a}$$

$$\frac{di_e(t)}{dt} = -\frac{i_e(t)}{\tau_{syn_e}} \tag{1b}$$

$$\frac{di_i(t)}{dt} = -\frac{i_i(t)}{\tau_{syn_i}} \tag{1c}$$

(a)   (b)   (c)   (d)   (e)   (f)
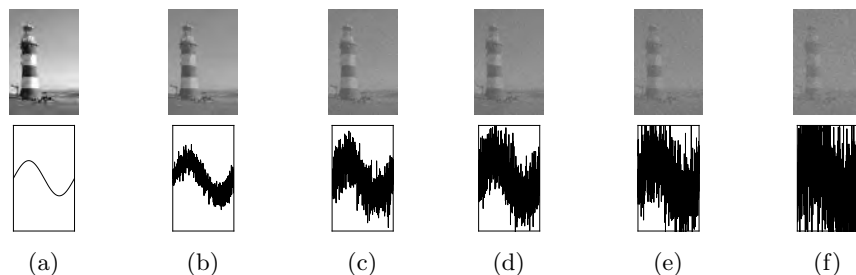
Figure 2: As an easy way to visualize the noise effects, the photographs (top row) had added to their greyscale values (0 to 255) noise proportional to how A_{noise} affects the membrane voltage, varying it from 0.0 (a) to 1.0 (f). On the bottom row, noise is applied to a sinusoid whilst keeping the same scale.

2.3 Benchmark Task

The benchmark test consisted of the simultaneous control of four joints (Figure 3) of a simulated BAXTER robot in order to draw a square shape on top of a table (for more details see [1]). All analyses are done on the robot's taskspace (Cartesian space) instead of joint space. Although being a two dimentional shape drawn on a surface, the system follows a human-inspired movement [4] and, for that reason, must keep in control a total of four dimensions: X, Y, Z and time.

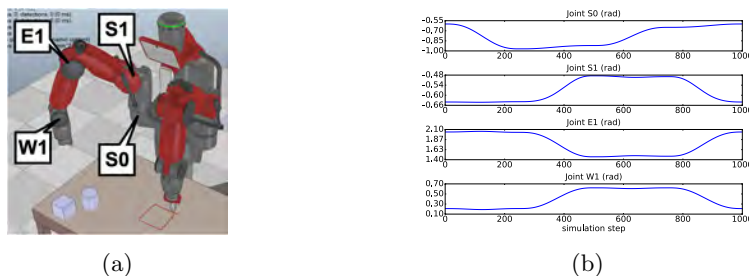(a)   (b)

Figure 3: Simulated BAXTER robot inside V-REP [17], with joint names indicated on its right arm, drawing the square on top of a table (a). Joint curves necessary to command the robot to generate the square shape (b).

Cost calculation Using the Dynamic Time Warping (DTW)[20] the total distance defined by the path formed with the minimum values of the accumulated distance (Figure 4, right-hand side) can be easily applied to compare the quality between different shapes. If the shapes are exactly the same, that distance is minimal and forms a straight diagonal line. The use of this same algorithm in a robotic task was already presented in [1].

3 Results and Discussion

Eleven distinct levels of noise were tested here for both, Modular and Monolithic, approaches (Section 2.1) with A_{noise} varying from 1.0 to 2.0 (Section 2.2). These experiments resulted in a total of $2,200$ simulations, where each one consisted of $3,000$ spiking neurons (five 600 neurons liquids in parallel). After every run, the joint values produced were loaded into the simulated Baxter robot inside V-REP to verify the final movement executed for the benchmark task and the results processed by the DTW algorithm (Section 2.3).

The DTW path cost values generated from three different noise intensities (A_{noise} equal to 1.0, 1.5 and 2.0) are presented in Figure 5 with all one hundred trials (bottom) and ten examples of the final shapes generated (top). Clearly as the noise is increased, the square shapes become strongly degraded,
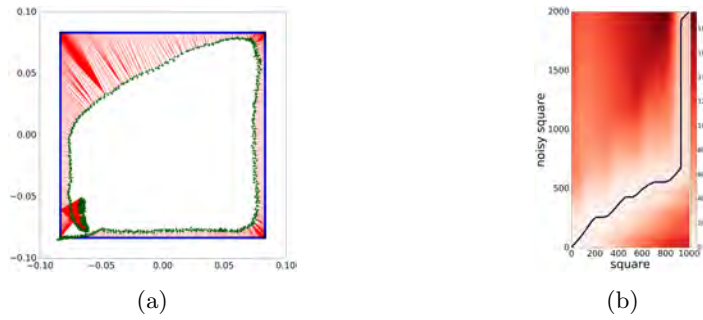
Figure 4: An example of how the DTW fits a very distorted square (green triangles) against a perfect one (blue squares) (a). The path that minimises the accumulated distance (b). The DTW path cost for the distorted square is 42.12.

(a) $A_{noise} = 1.0$

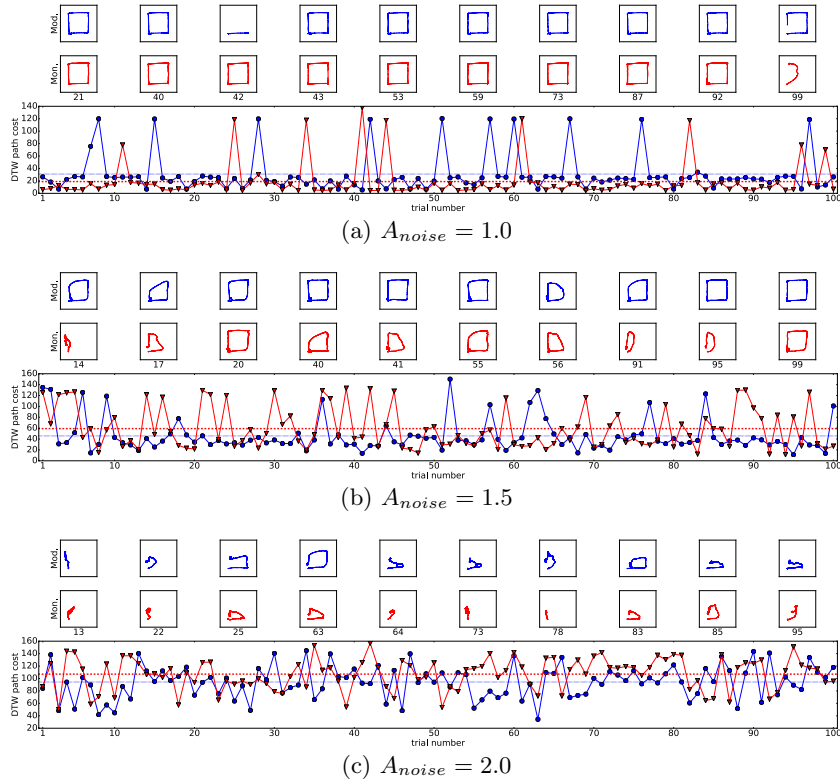(b) $A_{noise} = 1.5$

(c) $A_{noise} = 2.0$

Figure 5: Each plot shows the DTW path cost (bottom) for all trials and some of the shape outcomes (top) comparing the Modular (blue, circles) and the Monolithic approaches (red, triangles). The shapes (top) were selected based on the sorted cost values of both configurations to show a more comprehensive set of examples. Average values plotted as horizontal dashed lines (bottom).

but the Modular approach still can produce some rectangular forms even with $A_{noise} = 2.0$ or a noise level twice that injected during the readout training phase (see Figure 2 for a visual hint about noise levels). However, when using the standard noise level ($A_{noise} = 1.0$), the Monolithic approach had a better performance with an average cost value about 39% smaller than the Modular one. This type of system, sometimes, get stuck into a value and needs noise to be able to proceed, but the DTW algorithm penalises it as the trajectory it sees, although with a nice quality, was not completed. Therefore the difference between Modular and Monolithic approaches, with $A_{noise} = 1.0$, could be explained by the limited number of simulated steps (2,000 steps).
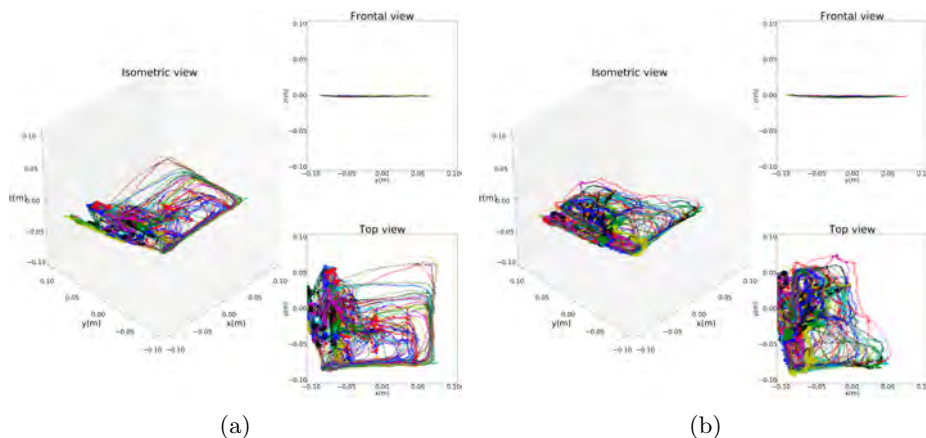

Figure 6: Modular (a) and Monolithic (b) approaches with A_{noise}=2.0.

In Figure 6, all hundred trials with $A_{noise} = 2.0$ were plotted together on 3D Cartesian space (same scale for all views) to make it easier to evaluate them, as mean values do not work well if there are time delays among trials. Despite the fact that a strong effect on the 2D square shape is clear, the Z axis (or the height control) is barely affected (top right).

The main question raised at the introduction was about the behaviour of this kind of system when affected by different noise levels and if it would have a nature like graceful degradation. To analyse that, the DTW path cost average and standard error values were calculated and are presented in Figure 7. The same figure also presents what would be the evolution of the cost considering the initial values incremented in steps of 10%.

Both approaches presented here, Modular and Monolithic parallel LSM, had what is considered a graceful degradation, as with the increase of the noise the systems did not catastrophically fail, but the DTW path cost grew in a well behaved manner. Comparing both LSM configurations, the Modular approach had an almost constant behaviour up to A_{noise}=1.4 when it started growing linearly with nearly the same slope as the Monolithic approach. Therefore, the
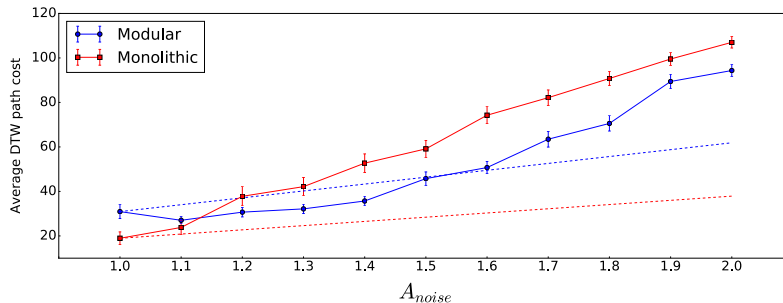
Figure 7: Average DTW path cost and its standard error for all trials (hundred in total for each A_{noise} level). The growth, considering the first value incremented by 10%, 20% ... 100%, is shown as a dashed line.

Modular system (between the A_{noise} range of 1.0 to 1.4) was able to withstand the noise better than a simple linear growth as showed by the dashed blue line (Figure 7) whilst the Monolithic configuration always increased its DTW path cost with the increase of noise.

4 Conclusions and Future Work

The robot controllers presented here were able to withstand, or at least gracefully degrade, when exposed to different noise levels - modelled here as white Gaussian noise based currents injected into the neuron model. These noisy currents could be seen, in a simplified way, as the consequences of exposition to non-destructive radiation. It is important to develop systems that are able to be implemented using new technologies, such as neuromorphic hardware, as they seem to be one of the possible ways to bypass the declining applicability of Moore's law [23] without having to expend huge amounts of energy [12]. Also, one of the strategies to decrease energy consumption, in a quadratic way, is the reduction of the voltage supplied to the digital circuits (near-threshold voltage [10]). However, this naturally leads to a decrease in the noise immunity as the voltage margin until a transistor changes its state is reduced. Another consequence of voltage reduction is within the speed a transistor changes its state. Still, neural systems are well known to be parallel, but relatively slow systems when compared to modern digital circuits. Even if MEMS-based logic gates [3] evolve up to the point of a final product, a digital system does not degrade gracefully in normal conditions and always needs extra gates to implement error correction.

The Modular design presented here opens up the possibility for a hot-swap hardware implementation, fitting SpiNNaker very well as it is able to turn on and off chips if necessary, and also decreasing the time and memory spent during learning. Also, having smaller readout layers, the time spent during learning is smaller than when using the Monolithic setup.

The Monolithic approach uses one big readout layer while the Modular one has smaller individual output layers and a node producing the average among them. In a future work, this simple average junction could be replaced by an extra on-line learning layer with weights connecting the analogue readout outputs directly to the neuron membrane, opening the possibility to vary the amount of trust the system has to each individual LSM without the need of changing the readout weights, thus saving energy and simplifying the design.

Additionally, to extend what was presented here, other parameters could be checked to verify their influence on the robustness. One good example, easily implemented, is the number of parallel liquids and the number of neurons used with each one.

In some trials, the systems got stuck in the middle of a well defined trajectory producing high DTW path cost values (see Figure 5a, trials 42 and 99). Our experience, after several experiments have been done using this type of system, together with the results presented in the Figure 7, suggests a certain minimum background noise is actually necessary for this kind of system. This idea of a "good" noise is not new [13] and will be left as another avenue for future works.

All the source code necessary to generate the results presented here will be available at http://github.com/ricardodeazambuja/ICONIP2016.

Acknowledgment

References

1. de Azambuja, R., Cangelosi, A., Adams, S.: Diverse, Noisy and Parallel: a New Spiking Neural Network Approach for Humanoid Robot Control. In: 2016 International Joint Conference on Neural Networks (IJCNN). p. In Press (In Press)
2. Benito-León, J., Louis, E.D.: Essential tremor: emerging views of a common disorder. Nature Clinical Practice Neurology 2(12), 666–678 (2006)
3. Chowdhury, F.K., Choe, D., Jevremovic, T., Tabib-Azar, M.: Design of MEMS based XOR and AND gates for rad-hard and very low power LSI mechanical processors. In: 2011 IEEE Sensors. pp. 762–765 (Oct 2011)
4. Flash, T., Hogan, N.: The coordination of arm movements: an experimentally confirmed mathematical model. The journal of Neuroscience 5(7), 1688–1703 (1985)
5. Furber, S.B., Lester, D.R., Plana, L.A., Garside, J.D., Painkras, E., Temple, S., Brown, A.D.: Overview of the SpiNNaker System Architecture. IEEE Transactions on Computers 62(12), 2454–2467 (Dec 2013)
6. Herculano-Houzel, S.: Scaling of Brain Metabolism with a Fixed Energy Budget per Neuron: Implications for Neuronal Activity, Plasticity and Evolution. PLoS ONE 6(3), e17514 (Mar 2011)

7. Indiveri, G., Linares-Barranco, B., Hamilton, T.J., van Schaik, A., Etienne-Cummings, R., Delbruck, T., Liu, S.C., Dudek, P., Häfliger, P., Renaud, S., Schemmel, J., Cauwenberghs, G., Arthur, J., Hynna, K., Folowosele, F., SAÏGHI, S., Serrano-Gotarredona, T., Wijekoon, J., Wang, Y., Boahen, K.: Neuromorphic silicon neuron circuits. Neuromorphic Engineering 5, 73 (2011)

8. Joshi, P., Maass, W.: Movement generation with circuits of spiking neurons. Neural Computation 17(8), 1715–1738 (2005)

9. Kandil, M.R., Tohamy, S.A., Abdel Fattah, M., Ahmed, H.N., Farwiez, H.M.: Prevalence of Chorea, Dystonia and Athetosis in Assiut, Egypt: A Clinical and Epidemiological Study. Neuroepidemiology 13(5), 202–210 (1994)

10. Kaul, H., Anders, M., Hsu, S., Agarwal, A., Krishnamurthy, R., Borkar, S.: Near-threshold voltage (NTV) design: opportunities and challenges. In: Proceedings of the 49th Annual Design Automation Conference. pp. 1153–1158. ACM (2012)

11. Kerns, S.E., Shafer, B.D., van Vonno, N., Barber, F.E.: The design of radiation-hardened ICs for space: A compendium of approaches. Proceedings of the IEEE 76(11), 1470–1509 (1988)

12. Kogge, P., Bergman, K., Borkar, S., Campbell, D., Carlson, W., Dally, W., Denneau, M., Franzon, P., Harrod, W., Hill, K., Hiller, J., Karp, S., Keckler, S., Klein, D., Lucas, R., Richards, M., Scarpelli, A., Scott, S., Snavely, A., Sterling, T., Williams, R.S., Yelick, K.: ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems. Tech. rep., University of Notre Dame (Sep 2008)

13. Maass, W.: Noise as a Resource for Computation and Learning in Networks of Spiking Neurons. Proceedings of the IEEE 102(5), 860–880 (May 2014)

14. Maass, W., Joshi, P., Sontag, E.D.: Computational Aspects of Feedback in Neural Circuits. PLoS Comput Biol 3(1), e165 (Jan 2007)

15. Maass, W., Natschläger, T., Markram, H.: Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural computation 14(11), 2531–2560 (Nov 2002)

16. Rogers, T.T., McClelland, J.L.: Parallel Distributed Processing at 25: Further Explorations in the Microstructure of Cognition. Cognitive Science 38(6), 1024–1077 (Aug 2014)

17. Rohmer, E., Singh, S.P., Freese, M.: V-REP: A versatile and scalable robot simulation framework. In: Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on. pp. 1321–1326. IEEE (2013)

18. Roy, D.S., Arons, A., Mitchell, T.I., Pignatelli, M., Ryan, T.J., Tonegawa, S.: Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. Nature 531(7595), 508–512 (Mar 2016)

19. Ryan, T.J., Roy, D.S., Pignatelli, M., Arons, A., Tonegawa, S.: Engram cells retain memory under retrograde amnesia. Science 348(6238), 1007–1013 (May 2015)

20. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE Transactions on Acoustics, Speech and Signal Processing 26(1), 43–49 (Feb 1978)

21. Schemmel, J., Brüderle, D., Grübl, A., Hock, M., Meier, K., Millner, S.: A Wafer-Scale Neuromorphic Hardware System for Large-Scale Neural Modeling. Proceedings of the 2010 IEEE International Symposium on Circuits and Systems (ISCAS"10) pp. 1947–1950 (2010)

22. Vertrees, S.M., Berman, S.A.: Chorea in Adults: Background, Pathophysiology, Epidemiology (Sep 2015), http://emedicine.medscape.com/article/1149854-overview, accessed: 2016-04-19

23. Waldrop, M.M.: More than Moore. Nature 530(7589) (Feb 2016)