Spacetime from Quantum Topology

Christopher Altman and Romàn R. Zapatrin

Abstract

Using adaptive quantum networks, we show how virtual spacetime structures emerge. Quantum parallelism, which is used to facilitate simultaneous training and revision of virtual network structures, gives rise to parallel evolution of linear superposed networks, providing optimal reconfiguration of whole-network quantum structure. The latter is interpreted under the model of virtual quantum spacetime. The main feature of this approach is that topology becomes a quantum value, subject to change under objective measurement.

1 Introduction

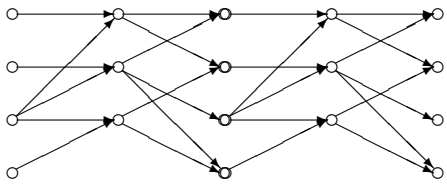
Over the last few decades, artificial neural networks (ANNs) have gained increasing interest, finding applications in various domains such as business planning, medicine, engineering, geology, and physics. ANNs have shown remarkable efficiency in solving tasks related to forecasting, planning, control and classification. ANNs are a powerful modeling tool, able to simulate complex dependencies. Due to their generic nonlinearity, the dimensionality — the required resources of the ANN — are smaller than those used in standard, linear optimization methods. The essential property of ANNs is their ability to be trained. Loosely speaking, instead of attempting to develop a theory, the network is trained on sample sets which it then learns to associate over a given domain or multimodal domains.

* † Faculty of Applied Physics, Kavli Institute of Nanoscience Delft, The Netherlands. Eamil: x@christopheraltman.com
Department of Informatics, The State Russian Museum, St.Petersburg, Russia. Email: roman.zapatrin@gmail.com

ANNs are a natural and intuitive field to explore, as they draw inspiration from the structure of real biological systems. One of the basic tasks of neural networks is to function as perceptrons, that is, to recognize signals for which we have no structural theory—for instance, to recognize visual patterns.

2 Training principles for ANNs

In this section, we review the basic principles on which perceptrons are based, their learning and training in classical setting, and show how Rota algebras—introduced above—emerge in their description. We shall deal with multilayered feedforward NNs, such as



that is, their nodes can be arranged in layers so that (i) no nodes in a given layer communicate, and (ii) the signals propagate only consecutively, via layers. The state of a given neural network is the set of values, numerical or vectorial, assigned to its nodes. A time-step propagation changes the perceptron's state so that the value of a node i at time $t+1$ depends on its own value and the values of nodes adjacent to i at time t.

Besides their multilayered, high-dimensional structure, the filtering and recognizing power of neural networks is based on the nonlinearity of their transition functions. In some cases, in particular while dealing with stability or elasticity issues, we may use linear approximations. This is exactly the case we explore in this paper: optimal topological configurations are assumed to be robust under small permutations [1].

Training. Initially, one starts with a set of patterns for which the classification is known. Usually by means of heuristic methods, the topological structure of the network is chosen and then trained via input of known patterns and subsequent adjustment of network parameters (transition functions). Output signals are correlated with patterns from different classes, to be well-separated with respect to certain criterion. The most popular method to adjust transition functions is error backpropagation (see, for instance,[2]).

Performance. Signal propagation in the linear approximation can be viewed as a matrix multiplication, which reduces to a number of arithmetic operations. The more links there are between neurons, the more computational resources are consumed by the process of pattern recognition. In order for a neural network to be faster, we should seek sparser configurations.

Therefore, the criterion for 'good matching' should also take performance into consideration. As an example we consider the approach of *fast neural networks*, which are weakly connected multilayered feedforward neural networks [1]. The idea of these networks grows from the FFT—the fast Fourier transform. When finding optimal configurations of such networks, we restrict the allowed links between neurons and force networks to be sparse.

3 From graphs to matrices

The core idea of the suggested approach is to convert the primary component of artificial neural networks – directed acyclic graphs (DAGs) – into a more general structure, which naturally complies with quantum parallelism. The latter inevitably means that this should be a linear structure in order to support superpositions. In [24] the idea to replace DAG by appropriate sets of linear matrices (associated further with appropriate quantum variables) was suggested.

The training paradigm remains the same as it was in conventional ANNs – standard performance criteria are used, such as Root Mean Square Error (RMSE) between targets and outputs of the neural network, the correlation (called ***R***-value) between the outputs and tar-gets an so on, the are only reclassified in terms of unitary matrices and matrix operations. The main peculiarity is that the result of training optimization is then a matrix, or a set of matrices, rather than a directed acyclic graph, as is seen in a standard, 'literal' approach.

We use the term *adaptive neural networks* for these structures. However, the primary objects subject to revision and training are operators in linear spaces – which can be realized, for instance, as quantum observables, rather than as explicitly defined neural networks. What do these matrices have to do with ANNs? In order to implement the result, that is, to recover the initial graph structure upon derivation of the appropriate solution, the Rota algebraic spatialization procedure [10] is applied.

Within our approach, the topology of a feedforward artificial neural network, $\mathcal{N}$, is described by the TEMPLATE MATRIX $\mathbf{A}$ of the appropriate directed acyclic graph (DAG), which is formed as follows:

$$\mathbf{A}_{jk} = \begin{cases} *, \text{ if } j \to k \text{ in } \mathcal{N} \\ 0, \text{ otherwise} \end{cases} \tag{1}$$

where $*$ stands for a wildcard – any number, and the set of such numerical matrices form an algebra [12] as it is closed under multiplication. The main property of $\mathbf{A}$ is that the synaptic weights 'follow' it – namely, if $\mathbf{A}_{jk} = 0$, then $w_{jk} = 0$. This differs from the standard description in terms of adjacency matrices, and the difference is that the matrix is not numeric. The simplest example is provided by a two-vertex DAG with one arrow

$\mathcal{N} =$  (2)

whose template matrix is

$$\mathbf{A} = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \tag{3}$$

In general, taking various matrices A by substituting the asterisk in template matrix ***etempl*** by numbers we may form various products ***esigprop***. The resulting set of matrices is called the Rota algebra of N . It can be verified that this set is closed under sums and matrix products, and thus qualifies as a closed algebra. This description is explicated under greater detail in [24].

4 Conventional training in linear terms

Given a feedforward neural network $\mathcal{N}$, consider the linear space $\mathcal{H}$, whose basis is labeled by the nodes of $\mathcal{N}$. A vector $\vec{x} \in \mathcal{H}$ is associated with a state of the network, namely the j-th component of $\vec{x}$ is the activity in node j. Note that the state of $\mathcal{N}$ captures the activities of *all* nodes, so a convention is required to define the default activity value of the nodes, we set its initial value to zero.

Initial activation associated with input vector $\mathbf{x}_0$ corresponds to changing values of only a part of nodes – specifically those in the input layer. The signal then propagates through $\mathcal{N}$, and we describe this activation by a linear operator W in $\mathcal{H}$. We start with $\mathbf{x}_0$, then:

$$\mathbf{x}_0 \;\mapsto\; \mathbf{x}_1 = \mathbf{x}_0 W \;\mapsto\; \ldots \;\mapsto\; \mathbf{x}_{\text{out}} \;=\; \mathbf{x}_n = \mathbf{x}_0 W^n \tag{4}$$

following n steps. Recall that postfix notation for operators acting on vectors is employed in this instance. The crucial feature of the suggested approach is that signal propagation is now described by iteration of *the same* operator W.

We emphasize that the operator W completely specifies the ANN in question, so from this step forward, the operator W, rather than the collection of weights, as in the traditional backpropagation approach, will be subject to training. A quantum analog of backpropagation training method has been described in [24]. We do not dwell on details, as the main point is that the result is a matrix, or a set of matrices, which may *not* be – as they are – associated with any DAG. We have no nodes, no arrows, just matrices. How to recover the spatial structure of the obtained optimal network? For that function, Rota's spatialization procedure is to be applied.

5 Rota topology: the emergence of spacetime structures

The matrix formalism deals with objects having no specific spatiotemporal features. We call them continuously-evolving, superposed topologies, but these topologies are yet 'spaceless' (while not 'timeless,' as we are speaking of their evolution in time, with respect to the proper time of the experimenter). The possibility to transfer them into spacetime manifold fluctuations was studied in quantum gravity.

The key point of the spatialization procedure is the following: consider the full matrix algebra $\mathcal{A}$ as a linear space. From this perspective, the Rota algebras Ω described by etempl are just linear subspaces of $\mathcal{A}$. Having a procedure which – starting from a subset of $\mathcal{A}$ – creates a topological space, providing the capability to discuss superposed configurations of differing neural networks. In this section we present a procedure which – starting from a given subspace of $\mathcal{A}$ – produces a set, and endows it with a topology that can be associated with certain acyclic directed graphs.

When the set of nodes is fixed, Rota algebras provide a suitable machinery to describe topology changes, which are expressed in terms of creation and annihilation of edges. In terms of template matrices etempl, these operations are adding or removing asterisks from the appropriate templates. In a fixed basis, a Rota algebra is a subspace of a full matrix algebra $\mathcal{A}$ – therefore adding or removing asterisks is equivalent to adding or removing basis vectors. Thus, any particular Rota algebra is a linear subspace of Ω. The basic idea of our

approach is to associate *any* subspace of Ω with an appropriate directed acyclic graph.

The first step of the spatialization procedure is then creating or finding points of the incidence algebra Ω, which will become nodes of the future graph:

$$\{\,\text{points}\,\} = \{\,\text{irreducible representations}\,\} \tag{5}$$

For commutative algebras this is a well-known Gel'fand transform, but in our case the algebra Ω is essentially non-commutative. We may, of course, consider a special case when Ω is commutative, but it will yield nothing but a trivial discrete topology on the set of the obtained points.

Rota topology. The primary function is to denote by $\mathbf{M}$ the set of points (that is, IRRs - irreducible representations) of Ω, each of which we shall associate with a prime ideal in Ω. Then there is an explicit algorithm, which associates a transitive graph with a subalgebra. The graph, in turn, can be treated as a topology on the set of its vertices (=points=IRRs). Note that in general, the Rota topology can be defined upon any set of ideals, and it is not necessary for Ω to be an algebra. Any linear subspace is suited for this purpose. The detailed description of the formalism can be found in [8].

To obtain an idea as to how the spatialization procedure works, it can be loosely described as finding the appropriate linear transformation, which renders the given set of matrices to that of upper triangular ones, which, in turn, resemble incidence matrices of appropriate DAGs. These graphs are the result of the spatialization.

6 Concluding remarks

As outlined in the mathematical formalism, it should be noted that *adaptive quantum networks*, as coherent quantum metastructures, are not standard neural networks. Rather, they are defined as superposed linear spaces, a meta-system whose subspaces consist of neural networks—just as a superposed wavefunction possesses an identity distinct from its constituent classical counterpart under influence of environmental interaction. In the ideal case, only following convergence to the predesignated, target network *minima* under backpropagation training, is projection via the spatialization procedure applied to convert the metasystem into a classical artificial neural network. As graphs are automatically produced as a consequence of this spatialization procedure, no *ab initio* association of states with graphs is requisite to implement the model into a candidate physical structure.

Coherent quantum information processing architectures both well-suited for implementation of adaptive quantum networks, and demonstrating promising experimental progress to date include highly-entangled, cluster state quantum computers suggested by Raussendorf and Briegel [13] and recent implementation by Zeilinger [14], high-dimensional Josephson junction qubit arrays [15], optical lattice traps in Bose-Einstein condensates [16], thin-layer diamond nanofilms [17], as well as novel materials currently under development by research teams active worldwide. Implementation of the training algorithm requires a relatively limited number of physical constraints to be satisfied—high-dimensional superposition, multiple degrees of internal freedom, and sufficient timescale under decoherence-free coherent state evolution. Consequently, adaptive quantum networks may be implemented both in state-of-the-art quantum technologies, as

well as unexpectedly demonstrated within naturally-occurring, dynamic molecular complexes, such as brain microtubules [20] or proteomics in molecular biology [21, 23]. Several natural systems have been observed to exhibit interference properties and high degrees of internal freedom – including fullerenes [19], tetraphenyl-porphyrin [19], small biomolecules [19], and chlorophyll [21]. Contrary to conventional expections, Briegel *et al.* recently demonstrated persistent, driven entanglement in non-equilibrium quantum systems coupled to a hot and noisy environment, in which the presence of environmental noise plays a constructive role in maintaining recurring quantum entanglement. As such, promising candidates in coherent quantum information processing technologies, condensed matter systems, and biophysics are all potential areas of investigation for experimental implementation of adaptive quantum networks.

Acknowledgements

We highly appreciate the friendly inspiration provided by Ignazio Licata.

References

[1] A. Dorogov, Structural Synthesis of Modular Weakly Connected Neural Networks. I. Methodology of Structural Synthesis of Modular Neural Networks, *Cybernetics and Systems Analysis* **37** 175-181 (2001)

[2] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, *Nature* **323** (1986) 533–536

[3] D. Ventura, T. Martinez, *An artificial neuron with quantum mechanical properties*, Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms 482–485 (1997).

[4] E. C. Behrman, J. E. Steck, S. R. Skinner, *A spatial quantum neural computer*, Proceedings of the International Joint Conference on Neural Networks, Washington D.C. (1999).

[5] R. L. Chrisley, *Quantum Computation and Quantum Learning*, Workshop on Brain, Mind and Physics, Prague (1993).

[6] D. Meyer, N. Wallach, Freedman, M. Z. Wang, *Topology and foundations of quantum algorithms*, US ARDA/NSA/ARO Quantum Computing Program Review, Nashville, TN (2003).

[7] C. Altman, J. Pykacz, R. R. Zapatrin, Superpositional Quantum Network Topologies, *International Journal of Theoretical Physics*, **43**, 2029 (2004).

[8] I. Raptis, R. R. Zapatrin, Algebraic description of spacetime foam, *Classical and Quantum Gravity*, **18**, 4187 (2001).

[9] P. J. Werbos, *The Roots of Backpropagation, From Ordered Derivatives to Neural Networks and Political Forecasting Adaptive and Learning Systems for Signal Processing, Communications, and Control Series*, John Wiley and Sons, (1994).

[10] R. R. Zapatrin, Finitary Algebraic Superspace, *International Journal of Theoretical Physics*, **37** 799–815 (1998).

[11] R.D. Sorkin, Indecomposable Ideals in Incidence Algebras, *Modern Physics Letters* **A18** 2491–2500 (2003); eprint math.CO/0309126.

[12] Rota, G.-C., *On The Foundation Of Combinatorial Theory, I. The Theory Of Möbius Functions,* Zeitschrift für Wahrscheinlichkeitstheorie, **2**, 340 (1968).

[13] R. Raussendorf, H. J. Briegel, *Physical Review Letters*, **86** 5188 (2001).

[14] M. S. Tame, R. Prevedel, M. Paternostro, P. Bohi, M. S. Kim and A. Zeilinger, Experimental Realization of Deutsch's Algorithm in a One-way Quantum Computer, *Physical Review Letters*, **98**, 140501 (2007)

[15] H. Mooij, The Road to Quantum Computing, *Science*, **307** 5713, 1210-1211 (2005).

[16] M. W. Zwierlein, C. A. Stan, C. H. Schunck, S. M. F. Raupach, S. Gupta, Z. Hadzibabic, and W. Ketterle, Observation of Bose-Einstein Condensation of Molecules, *Physical Review Letters* **91**: 250401 (2003).

[17] R. Hanson, D. Awschalom, Coherent manipulation of single spins in semiconductors, *Nature*, **453**, 1043–1049 (2008).

[18] A. Ferreira, A. Guerreiro, V. Vedral, Macroscopic Thermal Entanglement Due to Radiation Pressure, *Physical Review Letters*, **96**: 060407 (2006).

[19] M. Arndt, O. Nairz, J. Voss-Andreae, C. Keller, G. van der Zouw, A. Zeilinger, Wave-particle duality of C60, *Nature*, **401**, 680–682 (1999).

[20] S. Hagan, S.R. Hameroff and J.A. Tuszynski, Quantum Computation in Brain Microtubules: Decoherence and Biological Feasibility, *Physical Review* **E 65**, 61901:1-10 (2002)

[21] G. Engel, T. Calhoun, E. Read, T. Ahn, T. Manal, Y. Cheng, R. Blankenship, G. Fleming, Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems, *Nature*, **446**, 782–786 (2007).

[22] E. Nogales, D. Bouwmeester, Structure of the tubulin dimer by electron crystallography, *Nature*, **391**, 1099 (1998).

[23] J. Cai, S. Popsecu, H. J. Briegel, *Dynamic entanglement in oscillating molecules*, LANL technical report, quant-ph/0809.4906v1 (2008).

[24] Christopher Altman, Romàn R. Zapatrin, Backpropagation training in adaptive quantum networks, *International Journal of Theoretical Physics*, **49** 2991 (2010).