2 Peter 3:8 - Dominus videt in die, ut a mille annis, mille anni sicut dies.

Tua sublimitas tua asset, et minimum distantia nobis est curva sed recta linea.

HyperSpacetime: Complex Algebro-Geometric Analysis of Intelligence Quantum Entanglement Convergent Evolution

Yang Zhang Multiverse Corp hwswworld@yandex.com[a)]

(Dated: 7 April 2020)

Nature is structural instead of random, correlation is just approximation of causality, and data is not science: the more we reveal the more we revere nature on our voyage of unprecedented discovery. We argue that the soul(s) or exotic soul(s) of quotient Hypercomplex arbifold multiscale Spacetime (HyperSpacetime)'s corresponding manifold(s)/general (quotient and non-quotient) HyperSpacetime is the origin of super/general intelligence, and the metric of super/general intelligence is the complexity of quotient/general HyperSpacetime's corresponding generic polynomial. We also argue that the intersecting soul(s) and/or exotic soul(s) as varieties of quotient HyperSpacetime's corresponding manifold(s), when their maximal/minimum sectional curvatures approaching positive infinity and/or negative infinity as singularities, is the origin of quantum entanglement. We further argue that the maximal/minimum sectional curvatures of the same intersecting soul(s) and/or exotic soul(s), is the origin of convergent evolution through conformal transformation. We derive even N-dimensional HyperSpacetime, a M-open ($M = C^{I}_{I+N}, I, N, M \to \infty$) arbifold as generalized orbifold with the structure of a algebraic variety $\mathcal{A}$, without or with loop group action as $\mathcal{A} = [\mathcal{M}/\mathcal{LG}]$ ($\mathcal{M}$ as complex manifold, $\mathcal{LG}$ as loop group), it arises from I-degree (power of 2) hypercomplex even N-degree generic polynomial continuous/discrete function/functor as nonlinear action functional in hypercomplex $\mathbb{HC}^{\infty}$ useful for generic neural networks: $\mathcal{F}(S_j, T_j) = \prod_{n=1}^{N}(w_n S_n(T_n) + b_n + \gamma \sum_{k=1}^{j} \mathcal{F}(S_{k-1}, T_{k-1}))$ where $j = 1, \ldots, N$, $S_i = s_0 e_0 + \sum_{i=1}^{I-1} s_i e_i$, $T_i = t_0 e_0 + \sum_{i=1}^{I-1} t_i e_i$ over noncommutative nonassociative loop group. Its sectional curvature is $\kappa = \frac{|\mathcal{F}''(X)|}{(1+[\mathcal{F}'(X)]^2)^{\frac{3}{2}}}$ if $\mathcal{F}(X)$ is smooth, or $\kappa = \kappa_{max}\kappa_{min}$ if nonsmooth, by correlating general relativity with quantum mechanics via extension from 3+1 dimensional spacetime $\mathbb{R}^4$ to even N-dimensional HyperSpacetime $\mathbb{HC}^{\infty}$. By directly addressing multiscale, singularities, statefulness, nonlinearity instead of via activation function and backpropagation, HyperSpacetime with its corresponding generic polynomial determining the complexity of ANN, rigorously models curvature-based 2^{nd} order optimization in arbifold-equivalent neural networks beyond gradient-based 1^{st} order optimization in manifold-approximated adopted in AI. We establish HyperSpacetime generic equivalence theory by synthesizing Generalized Poincaré conjecture, soul theorem, Galois theory, Fermat's last theorem, Riemann hypothesis, Hodge conjecture, Euler's theorem, Euclid theorem and universal approximation theorem. Our theory qualitatively and quantitatively tackles the black box puzzle in AI, quantum entanglement and convergent evolution. Our future work includes HyperSpacetime refinement, complexity reduction and synthesis as our ongoing multiversal endeavor.

Keywords: Multiscale, Arbifold, Hypercomplex, Spacetime, Relativity, Entanglement, Evolution, Deep reinforcement learning, Complex analysis, Noncommutative, Nonassociative, Algebraic geometry, Geometric algebra, Singularity, Standard model, Wormhole, Exotic matter, Number theory, Orbifold, Manifold, Tensor, Gradient, Curvature, Polynomial, Loop group, Variety, Functor, PDE, Poincaré conjecture, Soul theorem, Universal approximation, Riemann hypothesis, Hodge conjecture, Galois theory, Fermat's theorem, Eulers theorem, Euclid theorem

I. INTRODUCTION

Far from the Madding Cloud

When groups of ants in haystack as agents
Descending along curvatures instead of gradients
Not powerful enough in overhauling environments
Evolving as adapters instead of gamers
Exploring and exploiting various actions
Searching and researching optimal states
Do they know they are on earth in universe multiverse

And the same as frogs in well hawks under sky
And dog's year different to human's year
And elephant in your eyes not the same as mine
Human in arbifolds just like ants in haystack
Hindsight insight foresight aspiration inspiration passion
Hate bacteria virus war all gone with wind
No hard times only great expectations

On the journey of human being's searching for beauty, simplicity and unification, especially in physics[132,133] and mathematics[13], from synthesis perspective, whether logic synthesis, physical synthesis, chemical synthesis, or biological synthesis, are all under the umbrella of commutative/noncommutative geometry[22,27,29], or Euclidean, Riemannian/elliptic/spherical and Lobachevsky/hyperbolic (same hyperbolic as Tanh[91]) geometries[65] as higher dimensional non Euclidean geometry with zero, positive and negative Gaussian/sectional curvature respectively, all supported by commutative/noncommutative algebra, associative/nonassociative algebra, such as Clifford algebra, tensor algebra[136], spin (Dirac, Pauli) algebra, and von Neumann algebra[23,49,67,127]. More specifically, universal geometry, quantum geometry[48], and biological geometry[103,140] like conformal geometry[42,94], are the outcome of physical laws and biological laws in modeling nonlinear physical and biological dynamics. Their significant applications adopting higher dimensional nonlinear manifold leveraging geometrization power frequently encountered in machine learning[84] such as linear regression[107], logistical regression, random forest, gradient boosting[74], Support vector machines (SVM)/kernel, decision trees, naive Bayes/prior, Nearest Neighbor, deep learning[68], whether supervised such as classification with labeled data, unsupervised such as clustering and Principal component analysis (PCA) with unlabeled data, self-supervised (context-based/temporal-based/contrasive-based), and semi-supervised like GAN[39] as well as reinforcement learning with no predefined data or even no data at all if with completely fixed rules, ranging from policy-based (deterministic/stochastic), value-based, model-based/model-free (trial-and-error), to actor-critic[73,120]. In most cases they may choose stochastic subgradient/gradient-descent or gradient-free approaches[2,12,30,31,50,70,141,142], or even Hessian matrix or Hessian free as well as orbifold with (as negative-curvature descent) or without adopting curvature-based approaches for higher dimensional unconstrained nonconvex optimization[3,8,71,77–79,82,87,89,124], all gear towards one single goal, that is, seeking causality from correlation through approximation. So far so good.

With radical paradigm shift and impressive progress in both hardware and software, now we can adopt various Artificial/Deep/Feedforward/Convolution/Recurrent (ANN/DNN/FNN/CNN/RNN) neural networks with billions of connections, billions of parameters, and hundreds of layers for real-life applications on facial recognition, speech recognition, language translation through universal function approximation, since the breakthrough made by AlexNet pioneered by LeNet and powered by GPU, along with its same all man-made successors including, AlexNet[7], VGGNet[113], GoogleNet[121], ResNet[46], DenseNet[57],

[a)] Multiverse Corp; Electronic mail: hwswworld@yandex.com

ShuffleNet[139], SqueezeNet[59], MobileNet[56], DeepComplexNet[123], DeepQuaternionNet[38], DeepOctonionNet[135], and NALU[125] on ImageNet benchmarks, as well as RNN/LSTM/seq2seq (end-to-end RNN)/RNN+attention (reinforcement learning/gradient descent)/transformer (full attention: non-self/self, hard-stochastic/soft-deterministic/multi-head, global/local)[51,53,63] based NLP applications such as GPT-2[101], BERT[40], ALBERT[66], Transformer-XL[26], XLNet[137], RoBERTa[72], CTRL[104], Megatron-LM[90]. However, the matter of fact is the degree of intelligence demonstrated by AI including deep learning[15] originated from perceptron, reinforcement learning, deep reinforcement learning,[83], AutoML[58] with or without[37] hyperparameter optimization, meta-learning, and neural architecture search (NAS)[143] and AutoDL all having exploration-exploitation trade-off dilemma, still falls far behind human intelligence in most cases. AI in adopting continuous optimization-centric gradient/subgradient-based deep reinforcement learning augmented with novel game theory such as mean field games, stochastic games, evolutionary games, beyond traditional and zero-sum game, and Convergent Evolution Strategies, as well as discrete optimization-centric gradient-free population-based genetic algorithms, has demonstrated awesome capability on beating human being in specific categories such as gaming. Since life is a game, so there is nothing wrong in tackling AI starting from gaming adopting deep reinforcement learning and evolution strategies as an alternative[105]: AlphaGo[111] AlphaZero,[112] DeepStack,[16] DeepCubeA,[116], and AlphaStar[5,129].

Despite of its wild success in certain domains, AI, particularly deep learning, has a few issues such as handcrafted neural network architecture, overwhelming number of weights, sensitivity of activation function[86,91], model size blowup, performance bottleneck, reproducibility crisis, and manual labor on labeling data for supervised learning and reward function design for reinforcement learning: imitation learning 100X more slower than human on learning how to drive, and much worse than that, reinforcement learning is 1,000X more slower than that of human. Even with the help of dedicated hardware like GPU and TPU, as well as better distributed processing middleware in favor of HPC-flavor MPI-type as opposed to cloud/data center-flavor RPC-type with shared memory burden and TCP/UDP overhead, there is still significant gap. We believe that the root cause of all of those problems is due to the black box nature of current AI practice. Even though there were a few attempts[108,110] in tackling the issue, however here we are dealing with more general open systems, hence conservation laws such as free-energy principle for closed systems do not apply any more.

II. A NEW PERSPECTIVE ON KNOWLEDGE

Human knowledge possesses a long and rich history: The very first sentence of Genesis 1:1, the Old Testament of the Holy Bible originally written roughly approximately in the 1660s B.C. to the 400s B.C., says, in the beginning God created the heaven and the earth. Sage Laozi, in his book titled Tao Te Ching written in the 600s B.C., says, Tao begets One (Taiji), One begets Two, Two begets Three, Three begets Everything. Coincidentally around the similar time frame, Greek philosopher Pythagoras believed multiverse (musica universalis) governed by mathematical equations, and metempsychosis (transmigration of souls) holding soul (gene in modern concept) being immortal as cycling among different living bodies as life.

Genes are DNA sequences that encode instructions for the synthesis of proteins, the total amount of DNA in a cell is referred to as genome, which lives in sequence space. There are approximately 60 trillion cells in human being. At any moment, human genome, which consists of 6.4 billion letters encoded in A (Adenine), C (Cytosine), G (Guanine), and T (Thymine), is being decoded to produce 20 possible amino acids for protein synthesis. Human genome contains the ensemble of the genetic heredity, and the instructions for both construction and operation. Through heredity, variations between individuals can accumulate and cause species to evolve As of today, the roles of most functional sequences in human genome still remain not completely decoded yet.

When Darwinism was developed, there was no concept of gene yet, and Darwinism addresses neither the source of life, nor the driver of life. Neo-Darwinism, also called the modern evolutionary synthesis, just like synthesis in EDA, generally denotes the integration of Charles Darwin's theory of evolution by natural selection, Gregor Mendel's theory of genetics by mutation as the basis for biological inheritance, and mathematical population genetics. General natural selection encompasses both external (environmentally driven) Darwinian natural selection, and internal self-organizing selection. Quantum Darwinism[144] even goes further, it aspires to explain the emergence of the classical world from quantum world as the superposition, entanglement, and environment as witness.

Knowledge as power has to be further powered by intelligence. Knowledge as gene of any living being, whether implicit or explicit, evolution-dependent or development-independent, agent-dependent or environment-dependent, formal such as DBpedia and Freebase[62] or informal, online (through web mining/data mining such as various search engines and domain specific apps, or crowd sourcing such as Wikipedia) or off-line, commonsense such as CYC[69] and Open Mind Common Sense[114,126] or niche, deduction-based or induction-based. Knowledge is not just power, it is mankind's gene passing from generation to generation. Knowledge is of paramount importance for any agent in dealing with uncertain environments without fixed rules, otherwise it is another story, for example, the success of AlphaGo and alike. Any living being who has genes, regardless with or without brain like virus and bacteria, possesses certain degree of intelligence. A lot of research on AI claims its inspiration comes from human brain, the fact is most of them are only related to visual cortex in cerebral instead. Brain, specifically cerebral cortex, used to be regarded as the sole origin of intelligence, unfortunately it is not. It has 16 billion neurons in contrast to cerebellum's 69 billion neurons. These neurons are connected to each other in a complex, recurrent fashion. It is still far away to completely understand how human brain exactly functions in demonstrating intelligence through brain science at present, and we don't know whether it is an open-end or dead-end.

Is human intelligence as well as other natural intelligence inborn by design via gene or developed through environment by learning? The answer is both, and that also applies to all living beings. According to one of latest research on human brain, human hippocampal neurogenesis drops sharply in children at age 13 to undetectable levels in adults, that explains exactly why children older than 13 cannot easily adapt to new languages. For example, epistemology is the branch of philosophy concerned with the theory of knowledge. Ontology in philosophy means the combination of subject and object, hence we can divide epistemology into intrinsic epistemology and extrinsic epistemology, as well as ontology into intrinsic ontology and extrinsic ontology. Both Intrinsic Epistemology and Intrinsic Ontology are pre-determined by gene as soul, and both Extrinsic Epistemology and Extrinsic Ontology are acquired after living being's birth by learning from environment. Being intrinsic means via heredity, evolution-dependent (agent-dependent) and development-independent, and being extrinsic means both environment-dependent and development-dependent. For HyperSpacetime based open systems, entropy cannot be conserved, hence it does not make any sense in discussing anything related with it such as information entropy.

III. SPACETIME AT A GLANCE

In physics, spacetime, Spacetime in recognizing the union of three spatial dimensions and one time dimension is well known for its adoption in general relativity,[33] which provides a unified description of gravity as a geometric property of continuous spacetime[45] at the largest scales such as universe, by generalizing special relativity and Newtonian theory. This is contrary to quantum mechanics[81] where spacetime is discrete and viewed as quantum many-body system. as a mathematical model, in recognizing the union of space and time, combines the three dimensions of space and the one dimension of time into a single 4-D continuum. Initially spacetime was proposed by Minkowski as a way to reformulate Einsteins special relativity (a special case of general relativity) right after its debut.

There are four types of fundamental interactions/forces in nature: gravity electromagnetism weak Interaction and strong Interaction. General relativity[34] leads to spectacular predictions as black holes, gravitational waves, and the big bang in early universe in macroscopic way, as what quantum mechanics does in microscopic way. Various efforts on developing Grand unified theory such as quantum electrodynamics (QED), quantum chrome-modynamics (QCD), and the standard model in unifying weak force, strong force and electromagnetism, and gravity, with the ultimate goal as unifying quantum mechincs and general relativity $QM = GR$[19,119], have been made, yet results are not perfect so far. In general relativity the gravitational field is encoded in spacetime as elliptic (Lorentzian) pseudo-Riemannian manifold. However, general relativity only models stand-alone systems, there are boundary-induced genuine spacetime singularities-triggered concerns at big bang and inside black holes[95]. Furthermore, when the curvature of spacetime becomes large enough on reaching the order of $\frac{1}{P_l^2}$, quantum effects[60] have to be taken into consideration as they start to dominates general relativity effects so that such curvature induced coordinate quantum singularities-triggered[64] concerns can be eliminated. At Planck scale, we must use an extended version of spacetime that fit for both general relativity and quantum physics. There are efforts on extending 3+1 dimensional spacetime model such as Kaluza-Klein model by introducing extra space dimension(s), but never goes to infinite or close to infinite space dimensions, and never extend time dimension beyond one, let alone allow time reversal. The traditional claims on the impossibility of going beyond 4-D spacetime is due to their linear partial differential equations (PDE) assumption while nature is inherently nonlinear[122]. A nonlinear dynamical system often can be described in nonlinear differential equations, such as Yang-Mills equation in quantum field theory, Boltzmann equation in statistical mechanics, Navier-Stokes equations in fluid dynamics, Lotka-Volterra equations in ecology, and Michaelis-Menten equations in enzyme kinetics. The hardness on solving those PDEs exactly in continuous optimization space, is similar to solve NP-complete and NP-hard problems in discrete optimization space. Amazingly as Planck scale is man-made, we can go even smaller physically by introducing HyperSpacetime model in higher dimensional non-Euclidean geometry beyond pseudo-manifold model in Riemannian geometry.

IV. GENERALIZING GENERAL RELATIVITY IN $\mathbb{HC}^\infty$

Based on general relativity, the Einstein field equations are formulated as follows and Wheeler[130] precisely summarize it as: matter tells spacetime how to curve, and spacetime tells matter how to move. In other words, gravity is geometry, matter sources gravity.

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi \frac{G}{C^4} T_{\mu\nu} \tag{1}$$

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \tag{2}$$

$\mu, \nu = 1, 2, 3, 4$ in elliptic (Lorentzian) pseudo-Riemannian manifold based 3+1 dimensional curved spacetime $\mathbb{R}^4$ on the action of Lie group SO(1,3) with metric signature (- + + +) adopting spacetime algebra[47], a Special Orthogonal (SO) finite dimensional Clifford algebra $Cl_{1,3}(R)$. Here $g_{\mu\nu}$ is metric tensor.

The initial value boundary problem in general relativity only gives us the metric on a patch of the spacetime. Other methods must be used to find the true global extension of that spacetime. Therefore, Einstein field equations alone cannot tell you the topology of the spacetime. Even ignore the above limitation, being nonlinear in nature, The general relativity Einstein field equations describe the relation between the geometry of a 3+1 dimensional elliptic (Lorentzian) pseudo-Riemannian manifold as Einstein manifold.

As nonlinear PDEs in modeling dynamical systems, Einstein field equations are very difficult to solve. de Sitter spacetime is a solution of the vacuum Einstein equations with a positive cosmological , and it is the maximally symmetric spacetime as elliptic (Lorentzian) pseudo-Riemannian manifold with positive curvature. Anti-de Sitter (Ads) spacetime[80] is a solution of the vacuum Einstein equations with a negative cosmological , and it is the maximally symmetric spacetime as hyperbolic (Lorentzian) pseudo-Riemannian manifold with negative curvature. There are other solutions, such as Schwarzschild solution, Reissner-Nordstrom solution, Kerr solution, and Friedmann solution.

General relativity and quantum cosmology are invariant under general spacetime diffeomorphisms (isomorphism of smooth manifolds). The quantum state of the universe is invariant under a time reversal change. The semi-classical state of the universe, has one definite direction of time (arrow of time). The processes occurring in the opposite direction of time seem to have disappeared in the actual universe. However, quantum entanglement may be telling us that they have not disappeared but they can be in a region of the spacetime that is not accessible for us. In fact, the time reversal invariance of the spacetime is broken in the semi-classical universe but a time symmetric solution always coexists because the time reversal invariance of the Friedmann equation. Therefore, if one consider that these two universes are created in entangled pairs, then, the time reversal symmetry does not disappear, it only lives in an inaccessible region.

The introduction of imaginary time[131] motivates us to extend (Lorentzian) pseudo-Riemannian manifold based 3+1 curved spacetime $\mathbb{R}^4$ to even N-dimensional curved HyperSpacetime in corresponding to M-open arbifold ($M = C^{I}_{I+N}$, $I, N, M \to \infty$) over $\mathbb{HC}^{\infty}$ in both imaginary time and imaginary space[61,102,106] over noncommutative nonassociative loop group constrained by multiscale[6,10,17,117] as opposed to planck scale.

$$g_{\mu\nu}(x) = \begin{bmatrix} g_{11}(x) & g_{12}(x) & g_{13}(x) & g_{14}(x) \\ g_{21}(x) & g_{22}(x) & g_{23}(x) & g_{24}(x) \\ g_{31}(x) & g_{32}(x) & g_{33}(x) & g_{34}(x) \\ g_{41}(x) & g_{42}(x) & g_{43}(x) & g_{44}(x) \end{bmatrix} \tag{3}$$

$\Longrightarrow$

$$g_{\alpha\beta}(X) = \begin{bmatrix} g_{11}(X) & g_{12}(X) & g_{13}(X) & g_{14}(X)\dots & g_{1N}(X) \\ g_{21}(X) & g_{22}(X) & g_{23}(X) & g_{24}(X)\dots & g_{2N}(X) \\ g_{31}(X) & g_{32}(X) & g_{33}(X) & g_{34}(X)\dots & g_{3N}(X) \\ g_{41}(X) & g_{42}(X) & g_{43}(X) & g_{44}(X)\dots & g_{4N}(X) \\ \dots & \dots & \dots & \dots & \dots \\ g_{N1}(X) & g_{N2}(X) & g_{N3}(X) & g_{N4}(X)\dots & g_{NN}(X) \end{bmatrix} \tag{4}$$

Here real number x becomes even I-degree hypercomplex number[11,106] X in hypercomplex number system, a generalization of complex numbers in higher dimension: $X = x_0 + \sum_{i=1}^{I-1}(x_i e_i)$

With that, in even N-dimensional HyperSpacetime $\mathbb{HC}^{\infty}$ ($M = C^{I}_{I+N}$, $I, N, M \to \infty$ and I as hypercomplex degree), the interval between two events ds can be defined as:

$ds^2 = c^2 \sum_{t=1}^{N}(X_{t2} - X_{t1})^2 - \sum_{s=1}^{N}(X_{s2} - X_{s1})^2$

and HyperSpacetime interval r can be defined as:

$r = \sqrt{c^2 \sum_{t=1}^{N}(X_{t2} - X_{t1})^2 + \sum_{s=1}^{N}(X_{s2} - X_{s1})^2}$

Operations on hypercomplex numbers, such as noncommutative quaternions, noncommutative plus nonassociative octonions as well as sedonions, and keep on going in 2^n as hypernions.

With that the Extended Einstein field equations are formulated as follows:

$$G_{\alpha\beta} + \Lambda g_{\alpha\beta} = 8\pi \frac{G}{C^4} T_{\alpha\beta} \tag{5}$$

$$G_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} \tag{6}$$

Where $\alpha, \beta = 1, 2, \dots, N$ all in hypercomplex $\mathbb{HC}^{\infty}$ over the action of noncommutative nonassociative loop group ($M = C^{I}_{I+N}$, $I, N, M \to \infty$).

V. EXTENDING QUANTUM MECHANICS IN $\mathbb{HC}^\infty$

Early quantum theory was profoundly re-conceived by Schródinger, Heisenberg, Born. There are two mathematical formalization for quantum mechanics which are equivalent: One is Heisenberg Picture, in which only the operators (observables and others) evolve in time, but the state vectors are constant with respect to time, an arbitrary fixed basis rigidly underlying the theory. The other is Schródinger Picture. in which only the state vectors evolve in time. but the operators (observables and others) are constant with respect to time. Dirac reconciliated the two pictures in Hilbert space and proved their equivalence[28] taking special relativistic effect into consideration. In classical mechanics observable (e.g. energy, position, momentum, etc.) is a function on a manifold called the phase space of the system. In contrary, quantum mechanical observable is an operator on a Hilbert space. Thus the commutative algebra of functions on it is replaced by the noncommutative algebra on a Hilbert space. Now it is von Neumann who gave the first complete mathematical formulation of this approach in terms of operators in Hilbert space, known as the Dirac-von Neumann axioms and von Neumann algebra. It is amazing from pure mathematical point of view, the infinite-dimensional state space in quantum mechanics offers a genuine multiverse/many-worlds interpretation of nature.

The Dirac equation was generalized to 3+1 dimensional curved spacetime[4] over $\mathbb{R}^4$ imposed by Planck scale, and like in what we do with spacetime in general relativity, we can easily further generalize it to even N-dimensional in corresponding to M-open arbifold $M = C^{I}_{I+N}$, $I, N, M \to \infty$ and I as hypercomplex degree. curved complex HyperSpacetime imposed by HyperSpacetime scale instead of Planck scale by making both imaginary time and imaginary space extensions as follows:

$$i\gamma^a e^\mu_a D_\mu \Psi - m\Psi = 0 \tag{7}$$

It is written by using Vierbein (frame) field/generalized Vierbein field, a set of 4 or N orthonormal vector fields interpreted as a model of 3+1 dimensional spacetime $\mathbb{R}^4$ or even N-dimensional HyperSpacetime $\mathbb{HC}^\infty$, and the gravitational spin connection. The Vierbein defines a local rest frame, allowing the $N * N$ as opposed to $4 * 4$ constant Dirac matrices γ^a to act at each spacetime point. Here $\mu = 1, \ldots, N$, $a = 1, \ldots, N$ both for even N-dimensional HyperSpacetime over $\mathbb{HC}^\infty$ as opposed to $\mu = 1, 2, 3, 4$, $a = 1, 2, 3, 4$ both for 3+1 dimensional spacetime over $\mathbb{R}^4$, e^μ_a is the Vierbein with $e^\mu_a e^\nu_a = g_{\mu\nu}$ as metric tensor in general relativity, and D_μ is the covariant derivative for fermionic fields defined as follows: $D_\mu = \partial_\mu - \dfrac{i}{2}\omega^{ab}_\mu \sigma_{ab}$ where σ_{ab} is the commutator of $N * N$ as opposed to $4 * 4$ Dirac matrices: $\sigma_{ab} = \frac{i}{2}[\gamma_a, \gamma_b]$ and ω^{ab}_μ are the spin connection components.

VI. HYPERSPACETIME GENERIC EQUIVALENCE

Axiom VI.1 *HyperSpacetime uncertainty axiom:*

$$\Delta x \Delta t > C_s \to 0 \tag{8}$$

with C_s as HyperSpacetime scale, a close to zero sub-Planck size, which makes Planck scale effect imposed by Heisenberg's Uncertainty principle irrelevant[18,32]. So does the string scale effect imposed in quantum gravity (Loops, M-theory including strings and branes where spacetime is not fundamental and time only has one-dimension[55,138]). The reason behind it is because the fine structure constant $\alpha = \frac{e^2}{\hbar c} \simeq \frac{1}{137}$ is still a mystery instead of a proved physical constant.

Proposition VI.1 *HyperSpacetime generic equivalence theorem: Even N-dimensional HyperSpacetime corresponding to M-open (connected, complete, noncompact, with no boundary) arbifold $\mathcal{A}$ ($M = C^{I}_{I+N}$, $I, N, M \to \infty$) with the structure of algebraic variety* *arising from I-degree (power of 2) hypercomplex even N-degree generic polynomial discrete/continuous function/functor $\mathcal{F}$ as nonlinear action functional in hypercomplex $\mathbb{HC}^\infty$:*

$$\mathcal{F}(S_j, T_j) = \prod_{n=1}^{N} (w_n S_n(T_n) + b_n + \gamma \sum_{k=1}^{j} \mathcal{F}(S_{k-1}, T_{k-1})) \tag{9}$$

where $j = 1, \ldots, N$, and hypercomplex number $S_i = s_0 e_0 + \sum_{i=1}^{I-1} s_i e_i$, $T_i = t_0 e_0 + \sum_{i=1}^{I-1} t_i e_i$. Its sectional curvature is $\kappa = \frac{|\mathcal{F}''(X)|}{(1+[\mathcal{F}'(X)]^2)^{\frac{3}{2}}}$ if $\mathcal{F}(X)$ is smooth, or as $\kappa = \kappa_{max}\kappa_{min}$ with $\kappa_\varphi = \kappa_{max}\cos^2\varphi + \kappa_{min}\sin^2\varphi$ where φ as the angle between the minimum principle plane with curvature κ_{max} and maximum principle plane with curvature κ_{min} if nonsmooth. Arbifold $\mathcal{A}$ can be Euclidean, spherical or hyperbolic depending on its underlying generic polynomial.

Arbifold $\mathcal{A}$ as M-open generalized orbifold is a quotient arbifold corresponding to quotient HyperSpacetime $\iff$ $\exists$ complex manifold(s) M and loop group(s) $\mathcal{LG}$, $\ni$ arbifold with loop group action as $\mathcal{A} = [\mathcal{M}/\mathcal{LG}]$. Otherwise Arbifold $\mathcal{A}$ is a non-quotient arbifold corresponding to non-quotient HyperSpacetime. Furthermore, $\mathcal{A}$ is a quotient arbifold $\iff$ $\mathcal{A}$'s corresponding manifold(s) M obey(s) Generalized Poincaré conjecture.

TABLE I. **HyperSpacetime Generic Polynomial's Applications on AI**

	$\mathcal{F}(S_j, T_j)$	S_i	T_i	w_n	b_n	γ **(Recurrent/Reinforcement)**
Deep Learning FNN	continuous	space	time	weight	bias	0
Deep Learning CNN	discrete	space	time	weight	bias	0
Deep Learning RNN	discrete	space	time	weight	bias	1
Reinforcement Learning	discrete/continuous	action	state	weight	bias	$0 < \gamma < 1$

The above proposition is synthesized from the established axiom and the following theorems, solved conjecture as lemmas as well as conjectures.

Lemma VI.2 *Universal approximation theorem*[25,54]*: Let $\varphi(.)$ be a nonconstant, bounded, and monotonically-increasing continuous function. Let I_m denote the m-dimensional unit hypercube $[0, 1]^m$. The space of continuous functions on I_{m0} is denoted by $C(I_m)$. Then, given any function $f \in C(I_m)$ and $\epsilon > 0$, there exist an integer N and sets of real constants α_i, $b_i \in R$, $w_i \in R^m$, where i = 1,...,N such that we may define:*

$F(x) = \sum_{i=1}^{N} \alpha_i \varphi\left(w_i^T x + b_i\right)$

as an approximate realization of the function f; that is, $|F(x)f(x)| < \epsilon$ for all $x \in I_m$.

Lemma VI.3 *Fermat's last theorem*[134]*: There are no positive integers x, y, z, and $N \geq 3$ such that $x^N + y^N = z^N$.*

Lemma VI.4 *Euler's theorem: Let φ be the angle, in the tangent plane, measured counterclockwise from the direction of minimum curvature κ_1 of minimum principle plane, and maximum curvature κ_2 of maximum principle plane. Then the normal curvature $\kappa_n(\varphi)$ in direction φ is given by $\kappa_n(\varphi) = \kappa_1\cos^2\varphi + \kappa_2\sin^2\varphi$.*

Lemma VI.5 *Euclid theorem: For any finite set of prime numbers, there exists a non-quotient number not in that set. In other words, there are infinitely many prime numbers, and there is no largest prime number.*

Lemma VI.6 *Fundamental theorem of Galois theory*[85]*: Let L/K be a finite Galois extension. Let $Gal(L/K)$ denote the Galois group of the extension L/K. Let H denote a subgroup of $Gal(L/K)$ and F denote an intermediate field. The mappings: $H \longmapsto LH$, and $F \longmapsto Gal(L/F)$ are inclusion-reversing and inverses. Moreover, these maps induce a bijection between the normal subgroups of $Gal(L/K)$ and the normal, intermediate extensions of L/K. .*

The above lemma's application on solutions of rational polynomial equation as follows: A polynomial equation is solvable by radicals $\iff$ its underlying Galois group is a solvable group. Hence for polynomial equations with degree $N > 4$, they are not solvable by radicals as there is no underlying Galois group being a solvable group.

Conjecture VI.7 *Hodge conjecture[52]: On a projective non-singular algebraic variety over $\mathbb{C}^N$, any Hodge class is a rational linear combination of classes $cl(Z)$ of algebraic cycles.*

Conjecture VI.8 *Riemann hypothesis: The Riemann zeta-function $\zeta(s)$ is a function of a complex variable defined by : $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ using analytical continuation for all complex $s \neq 1$, and all of the nontrivial zeroes of this function $\zeta(s) = 0$ lie on a vertical straight line with real part equal to exactly $1/2$.*

Lemma VI.9 *Generalized Poincaré conjecture (solved)[36,97–99,115] Every homotopy sphere (a closed N-manifold which is homotopy equivalent to the N-sphere) in the chosen category, i.e. topological manifolds, piecewise linear manifolds, or differential manifolds, is isomorphic to the standard N-sphere. The above claim is true in all dimensions for topological manifolds; true in dimensions other than 4; unknown in 4 for piecewise linear manifolds; false generally, true in some dimensions including 1,2,3,5, and 6, unsettled in 4 for differential manifolds.*

Proposition VI.10 *Generalized soul proposition: Suppose that (M, g) is an open (connected, complete, noncompact, with no boundary) manifold M with nonnegative (when M being spherical) or nonpositive (when M being hyperbolic) sectional curvature g, then M contains a soul (when M being spherical) or exotic soul (when M being hyperbolic) $S \subset M$, which is a compact, totally geodesic, totally convex (when M being spherical) or nonconvex (when M being hyperbolic) subarbifold; otherwise M contains no soul. Furthermore, M is diffeomorphic (when M being spherical) or diffeomorphic when M being hyperbolic) to the total space of the normal bundle of the S in M.*

The above proposition is synthesized from the following Soul theorem[21,41,44,96] as lemma and Exotic soul proposition:

Proposition VI.11 *Exotic soul proposition: Suppose that (M, g) is an open (connected, complete, noncompact, with no boundary) Riemannian manifold M of nonpositive sectional curvature g, then M contains an exotic soul $S_a \subset M$, which is a compact, totally geodesic, totally nonconvex submanifold. Furthermore, M is diffeomorphic to the total space of the normal bundle of the S_a in M. If (M, g) has negative sectional curvature, then any exotic soul of M is a point, and consequently M is diffeomorphic to $\mathbb{R}^N$.*

The above proposition is synthesized from the following Soul theorem as lemma:

Lemma VI.12 *Soul theorem: Suppose that (M, g) is an open (connected, complete, noncompact, with no boundary) Riemannian manifold M of nonnegative sectional curvature g, then M contains a soul $S \subset M$, which is a compact, totally geodesic, totally convex submanifold. Furthermore, M is diffeomorphic to the total space of the normal bundle of the S in M. If (M, g) has positive sectional curvature, then any soul of M is a point, and consequently M is diffeomorphic to $\mathbb{R}^N$.*

Proposition VI.13 *Metric of super/general intelligence proposition: The complexity of quotient/general HyperSpacetime's corresponding generic polynomial is the metric of super/general intelligence.*

Conjecture VI.14 *Origin of super/general intelligence conjecture: The soul(s) or exotic soul(s) of quotient/general HyperSpacetime's corresponding manifold(s) is the origin of super/general intelligence.*

Conjecture VI.15 *Origin of quantum entanglement conjecture: The intersecting soul(s) and/or exotic soul(s) as varieties of quotient HyperSpacetime's corresponding manifold(s), when their maximum sectional curvatures approaching positive infinity and/or negative infinity as singularities, is the origin of quantum entanglement*[1,9,14,20,24,35,43,75,76,88,92,93,100,109,118,128].

The above conjecture is partially generalized from $ER = EPR$ conjecture on the possibility of bridging EPR quantum entanglement as black holes and ER bridge as wormholes.

Conjecture VI.16 *Origin of convergent evolution conjecture: The maximum sectional curvatures of intersecting soul(s) and/or exotic soul(s) as algebraic varieties in quotient HyperSpacetime's corresponding manifold(s), is the origin of convergent evolution.*

TABLE II. **Dimension of I-degree Polynomial N-degree HyperSpacetime M-Abifold**

$M = C^I_{I+N} = \frac{(I+N)!}{I!N!}$	$N = 6$	$N = 8$	$N = 10$	$N = 12$	$N = 14$	$N = 16$
I=2(Complex)	28	45	66	91	120	153
I=4(Quaternion)	210	495	1001	1820	3060	4845
I=8(Octonion)	3003	12870	43758	125970	319770	735471
I=16(Sedonion)	74613	735,471	5311735	30421744	145422675	601080390

TABLE III. **Measuring General Intelligence of CNNs**

CNN	**Complexity of Underlying Equivalent Generic Polynomial**
AlexNet	left for its authors
VGGNet	left for its authors
GoogleNet	left for its authors
ResNet	left for its authors
DenseNet	left for its authors
ShuffleNet	left for its authors
SqueezeNet	left for its authors
MobileNet	left for its authors
DeepComplexNet	left for its authors
DeepQuaternionNet	left for its authors
DeepOctonionNet	left for its authors

TABLE IV. **Measuring General Intelligence of Transformers**

Transformer	**Complexity of Underlying Equivalent Generic Polynomial**
GPT-2	left for its authors
BERT	left for its authors
ALBERT	left for its authors
Transformer-XL	left for its authors
XLNet	left for its authors
RoBERTa	left for its authors
CTRL	left for its authors
Megatron-LM	left for its authors

VII. ACKNOWLEDGEMENTS

In case when citing our work in your paper, please use the following BibTeX entry:
@misc{CMOSpacetime,
author = {Yang Zhang},
Title = {COMSpacetime: Geometric-Algebraic Complex Analysis of Intelligence/Quantum

Entanglement/Convergent Evolution},
Year = {2020},
url = {https://www.vixra.org/pdf/2003.0206v1.pdf}}
@misc{MSpacetime,
author = {Yang Zhang},
Title = {Multiversal SpaceTime (MSpaceTime) Not Neural Network as Source of Intelligence in Generalized Quantum Mechanics, Extended General Relativity, Darwin Dynamics for Artificial Super Intelligence Synthesis},
Year = {2019},
url = {https://www.biorxiv.org/content/10.1101/858423v1.full.pdf}}
@misc{ArtificialDesign,
author = {Yang Zhang},
Title = {Artificial Design: Modeling Artificial Super Intelligence with Extended General Relativity and Universal Darwinism via Geometrization for Universal Design Automation},
Year = {2019},
url = {https://openreview.net/pdf?id=SyxQ_TEFwS}}
@misc{HyperSpacetime,
author = {Yang Zhang},
Title = {HyperSpacetime: Complex Algebro-Geometric Analysis of Intelligence Quantum Entanglement Convergent Evolution },
Year = {2020}}

[1]Scott Aaronson, Nai-Hui Chia, Han-Hsuan Lin, and Ruizhe Zhang Chunhao Wang. On the quantum complexity of closest pair and related problems, 2019.

[2]A. Anandkumar. Recent advances and challenges in non-convex optimization. In *ICML Workshop on Adaptive and Multitask Learning*, 2016.

[3]Rohan Anil, Vineet Gupta, Tomer Koren, Kevin Regan, and Yoram Singer. Second order optimization made practical, 2020.

[4]Mayeul Arminjon and Frank Reifler. Equivalent forms of dirac equations in curved spacetimes and generalized de broglie relations, 2013.

[5]Kai Arulkumaran, Antoine Cully, and Julian Togelius. An evolutionary computation perspective, 2019.

[6]C. Auffray and D. Noble. Scale relativity: an extended paradigm for physics and biology? *Foundations of Science*, 16(4):303–305.

[7]Ivana Balarzevic, Carl Allen, and Timothy Hospedales. Imagenet classification with deep convolution neural networks. In *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc., 2012.

[8]Ivana Balarzevic, Carl Allen, and Timothy Hospedales. Multi-relational poincaré graph embeddings. In *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019.

[9]Vijay Balasubramanian, Asad Naqvi, and Joan Simon. A multi-boundary ads orbifold and dlcq holography: A universal holographic description of extremal black hole horizons, 2004.

[10]Itzhak Bars. Gauged duality, conformal symmetry and space-time with two times. *Phys. Rev.*, D58, 1998.

[11]Carl M. Bender, Dorje C. Brody, and Daniel W. Hook. Quantum effects in classical systems having complex energy, 2008.

[12]Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives, 2014.

[13]J. Bernstein and S. Gelbart. *An introduction to the Langlands Program*. Birkhauser, Boston, 2003.

[14]W. P. Bowen and P. K. Lam. Biased epr entanglement and its application to teleportation, 2002.

[15]Michael M. Bronstein. Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, 34:18–42, 2017.

[16]N. Brown and T. Sandholm. Safe and nested subgame solving for imperfect-information games. In *Neural Information Processing Systems (NIPS)*. NIPS Foundation, 2017.

[17]Gianluca Calcagni. Multiscale spacetimes from first principles. *Physical Review D*, 95(064057), 2017.

[18]X. Calmet, M. Graesser, and SD Hsu. Minimum length from quantum mecahnics and general relativity. *Physics Review Letter*, 93(211101), 2004.

[19]ChunJun Cao and Sean M. Carroll. Bulk entanglement gravity without a boundary: Towards finding einstein's equation in hilbert space, 2017.

[20]Carlos Castro. Noncommutative geometry, negative probabilities and cantorian-fractal spacetime, 2000.

[21]Jeff Cheeger and Detlef Gromoll. On the structure of complete manifolds of nonnegative curvature. *The Annals of Mathematics*, 96(2):413–443, 1972.

[22]Alain Connes. *Noncommutative Geometry*. Academic Press, 1994.

[23]Alain CONNES and Michel DUBOIS-VIOLETTE. Yang-mills and some related algebras, 2004.

[24]Christophe Couteau. Spontaneous parametric down-conversion, 2018.
[25]G. Cybenko. Approximations by superpositions of sigmoidal functions. *Mathematics of Control, Signals, and Systems*, 2(4):303–314, 1989.
[26]Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc V. Le, and Ruslan Salakhutdinov. Transformer-xl: Attentive language models beyond a fixed-length context, 2019.
[27]Robbert Dijkgraaf. The mathematics of m-theory. *European Congress of Mathematics*, 201:1–19, 2001.
[28]Paul A. M. Dirac. *The Principles of Quantum Mechanics.* Oxford University Press, Oxford, 1930.
[29]Simon K. Donaldson. Yang-mills theory and geometry. URL `http://wwwf.imperial.ac.uk/~skdona/YMILLS.PDF`.
[30]Xiao Dong and Ling Zhou. Understanding over-parameterized deep networks by geometrization, 2019.
[31]Xiao Dong and Ling Zhou. Geometrization of deep networks for the interpretability of deep learning systems, 2019.
[32]M. J. Duff. Comment on time-variation of fundamental constants, 2002.
[33]Albert Einstein. The albert einstein archives. URL `http://www.albert-einstein.org`.
[34]Albert Einstein. *The Meaning of Relativity.* Princeton University Press, 1923.
[35]Lukasz Fidkowski, Veronika Hubeny, Matthew Kleban, and Stephen Shenker. The black hole singularity in ads/cft, 2003.
[36]Michael Freeman. The topology of four-dimensional manifolds. *J. Differential Geom.*, 17(3):357–453, 1982.
[37]Adam Gaier and David Ha. Weight agnostic neurual networks, 2019.
[38]Chase Gaudet and Anthony Maida. Deep quaternion networks, 2018.
[39]Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. In *Neural Information Processing Systems (NIPS).* NIPS Foundation, 2014.
[40]Google. URL `blog.google/products/search/search-language-understanding-bert`.
[41]Detlef Gromoll and Wolfgang Meyer. On complete open manifolds of positive curvature. *The Annals of Mathematics*, 90(2):75–90, 1969.
[42]X. Gu, Y. Wang, T. F. Chan, P. M. Thompson, and S.-T. Yau. Genus zero surface conformal mapping and its application to brain surface mapping. In *Information Processing in Medical Imaging,18th International Conference*, pages 172–184, 2003.
[43]Eduardo Guendelman, Emil Nissimov, Svetlana Pacheva, and Michail Stoilov. Einstein-rosen bridge revisited and lightlike thin-shell wormholes. In *3rd National Congress on Physical Sciences*, 2016.
[44]Luis Guijarro and Gerard Walschap. The curvature operator at the soul, 2011.
[45]S. W. Hawking and G. Ellis. *The Large Scale Structure of Space-Time.* Cambrdige University Press, 1973.
[46]Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016.
[47]D. Hestenes. Spacetime physics with geometric algebra. *American Journal of Physics with geometric algebra*, 71(6):691, 2003.
[48]Hoshang Heydari. Geometric formulation of quantum mechanics, 2017.
[49]Eckhard Hitzer, Tohru Nitta, and Yasuaki Kuroe. Applications of clifford's geometric algebra, 2013.
[50]S-S Ho, P Dai, and F Rudzicz. Manifold learning for embedding multivariate variable-length sequences with an application to similarity search. *IEEE Transactions on Neural Networks and Learning Systems*, 27(6):1333–1344, 2015.
[51]S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computing*, (9):1735–1780, 1997.
[52]W. V. D. Hodge. The topological invariants of algebraic varieties. In *Proceedings of the International Congress of Mathematicians*, pages 181–192, 1950.
[53]J. J. Hopfield. Neural network and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sc.*, 79(8):2554–2558, 1982.
[54]K. Hornik, M. Stinchcombe, and H. Whit. Multilayer feedforward networks are universal approximators. *Journal Neural Networks*, 2(5):359–366, 1989.
[55]Gary T. Horowitz. Spacetime in string theory, 2004.
[56]Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017.
[57]Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks. In *CVPR*, 2017.
[58]Frank Hutter, Lars Lottoff, and Joaquin Vanschoren (Eds.). *AutoML: Methods, Systems, Challenges.* Springer-Verlag, 2019.
[59]Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, and Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and ¡0.5mb model size, 2016.
[60]Shing-Tung Yau Juven Wang, Xiao-Gang Wen. Quantum statistics and spacetime topology: Quantum surgery formulas, 2020.
[61]Gerald Kaiser. Quantum physics, relativity, and complex spacetime towards a new synthesis, 2009.
[62]Mayank Kejriwal and Daniel P. Miranker. Experience: Type alignment on dbpedia and freebase, 2016.
[63]Boris Knyazev, Graham Taylor, and Mohamed Amer. Understanding attention in graph neural networks. In *International Conference on Learning Representations (ICLR) Workshop on Representation Learning*

on Graphs and Manifolds, 2019. URL `http://arxiv.org/abs/1905.02850`.

[64] Deborah Konkowski and T. M. Helliwell. Quantum singularities in static and conformally static spacetimes. *International Journal of Modern Physics*, 26(22), 2011.

[65] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marian Boguna. Hyperbolic geometry of complex networks, 2010.

[66] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut. Albert: A lite bert for self-supervised learning of language representations, 2020.

[67] Sylvain Lavau. Tensor hierarchies and leibniz algebras, 2019.

[68] Y. LeCun, Y. Bengio., and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[69] D. B. Lenat. Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38, 1995.

[70] Mario Lezcano-Casado and David Martinez-Rubio. Cheap orthogonal constraints in neural networks: A simple parametrization of the orthogonal and unitary group, 2019.

[71] Mingrui Liu and Tianbao Yang. On noisy negative curvature descent: Competing with gradient descent for faster non-convex optimization, 2017.

[72] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach, 2019.

[73] Yuxi Liu. Deep reinforcement learning, 2018.

[74] F. Locatello, G. Dresdner, I. Valera, and G. Ratsch. Boosting black box variational inference. In *NIPS*, 2018.

[75] Juan Maldacena. Eternal black holes in ads, 2001.

[76] Juan Maldacena and Leonard Susskind. Cool horizons for entangled black holes, 2013.

[77] James Martens. Deep learning via hessian-free optimization. In *ICML 2010*, 2010.

[78] James Martens. New insights and perspectives on the natural gradient method, 2017.

[79] James Martens, Jimmy Ba, and Matt Johnson. Kronecker-factored curvature approximations for recurrent neural networks. In *International Conference on Learning Representations (ICLR)*, 2018.

[80] Brett McInnes. Orbifold physics and de sitter spacetime, 2004.

[81] Jagdish Mehra. *The Historical Development of Quantum Theory 1-6*. Springer-Verlag, New York, 2001.

[82] Marko Valentin Micic and Hugo Chu. Hyperbolic deep learning for chinese natural language understanding, 2018.

[83] Sobhan Miryoosefi, Kianté Brantley, Hal Daumé III, Miroslav Dudík, and Robert Schapire. Reinforcement learning with convex constraints. In *NeurIPS*, 2019.

[84] T. M. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.

[85] Shinichi Mochizuki. Galois sections in absolute anabelian geometry, 2014.

[86] Alejandro Molina, Patrick Schramowski, and Kristian Kersting. Pade activation units: End-to-end learning of flexible activation functions in deep networks, 2020.

[87] Andreas Mutter, Erik Parr, and Patrick K.S. Vaudrevange. Deep learning in the heterotic orbifold landscape, 2019.

[88] M.S. El Naschie. On the uncertainty of cantorian geometry and the two-slit experiment. *Chaos Solitons and Fractals*, 19(3):517–529, 1998.

[89] Maximilian Nickel and Douwe Kiela. Poincaré embeddings for learning hierarchical representations. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 6341–6350. 2017.

[90] Nvidia. URL `github.com/NVIDIA/Megatron-LM`.

[91] Chigozie Nwankpa, Winifred Ijomah, Anthony Gachagan, and Stephen Marshall. Activation functions: Comparison of trends in practice and research for deep learning, 2018.

[92] C. F. Ockeloen-Korppi, E. Damskagg, J.-M. Pirkkalainen, M. Asjad, A. A. Clerk, F. Massel, M. J. Woolley, and M. A. Sillanpaa. Stabilized entanglement of massive mechanical oscillators. *Nature*, 556: 478–482, 2018.

[93] Daniele Oriti. Spacetime as quantum many-body system, 2017.

[94] Andras Pellionisz and Rodolfo Llinas. Tensor theory of brain function. the cerebellum as a space-time metric,. In *Competition and Cooperation in Neural Nets*, pages 394–117, 1982.

[95] Roger Penrose. Conformal treatment of infinity. *General Relativity and Gravitation*, 43:901–922, 2011.

[96] Grigori Perelman. Proof of the soul conjecture of cheeger and gromoll. *On complete open manifolds of positive curvature*, 40(1):209–212, 1994.

[97] Grigori Perelman. The entropy formula for the ricci flow and its geometric applications, 2002.

[98] Grigori Perelman. Finite extinction time for the solutions to the ricci flow on certain three-manifolds, 2003.

[99] Grigori Perelman. Ricci flow with surgery on three-manifolds, 2003.

[100] Mark Van Raamsdonk. Building up spacetime with quantum entanglement. *General Relativity and Gravitation*, 42(10):2323–2329, 2010.

[101] Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. 2019.

[102] Del Rajan. Complex spacetimes and the newman-janis trick, 2017.

[103] N. Rashevsky. The geometrization of biology. *The bulletin of mathematical biophysics*, 18(1):31–56, 1956.

[104]SalesForce. URL `einstein.ai/presentations/ctrl.pdf`.
[105]Tim Salimans. Evolution strategies as a scalable alternative to reinforcement learning, 2017.
[106]Abdul Rouf Samurah. M-theory on complex spacetime, 2018.
[107]Craig Saunders, Alexander Gammerman, and Vladimir Vovk. Ridge regression learning algorithm in dual variables. In *ICML*, 1998.
[108]Andrew M. Saxe, Yamini Bansal, Joel Dapello, and Madhu Advan. On the information bottleneck theory of deep learning. In *International Conference on Learning Representations (ICLR)*, 2018.
[109]Nathan Seiberg and Edward Witten. The d1/d5 system and singular cft, 1999.
[110]Ravid Shwartz-Ziv and Naftali Tishby. Opening the black box of deep neural networks via information, 2017.
[111]David Silver. Mastering the game of go with deep neural network and tree search. *Nature*, 529(7587): 484–489, 2016.
[112]David Silver. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.
[113]K Simonyan and A Zisserman. Very deep convolutional networks for large-scale image recognition. 2014.
[114]Push Singh, Thomas Lin, Erik T. Mueller, Grace Lim, Travell Perkins, and Wan Li Zhu. Open mind common sense: Knowledge acquisition from the general public. In *On the Move to Meaningful Internet Systems 2002*, pages 1223–1237, 2000.
[115]Stephen Smale. Generalized poincaré's conjecture in dimensions greater than 4. *The Annals of Mathematics*, 74(2), 1961.
[116]Pierre Baldi Stephen McAleer, Alexander Shmakov. Solving the rubik's cube with deep reinforcement learning and search. *Nature Machine Intelligence*, 1:356–363, 2019.
[117]Luka Stopar, Primoz Skraba, Marko Grobelnik, and Dunja Mladenic. Streamstory: Exploring multivariate time series on multiple scales. *IEEE Transactions on Visualization and Computer Graphics*, 99 (1-1), 2018.
[118]Leonard Susskind. Er=epr, ghz, and the consistency of quantum measurements, 2014.
[119]Leonard Susskind. Dear qubitzers, gr=qm, 2017.
[120]Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction, Second Edition*. MIT Press, Cambridge, MA, 2018.
[121]Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *CVPR*, 2015.
[122]M. Tegmark. On the dimensionality of spacetime. *Classical and Quantum Gravity*, 1997.
[123]Chiheb Trabelsi, Olexa Bilaniuk, Ying Zhang, Dmitriy Serdyuk, Sandeep Subramanian, Joao Felipe Santos, Soroush Mehri, Negar Rostamzadeh, Yoshua Bengio, and Christopher J Pal. Deep complex networks, 2018.
[124]Lucas Vinh Tran, Yi Tay, Shuai Zhang, Gao Cong, and Xiaoli Li. Hyperml: A boosting metric learning approach in hyperbolic space for recommender systems. In *WSDM20: Proceedings of the 13th International Conference on Web Search and Data MiningJanuary 2020*.
[125]Andrew Trask, Felix Hill, Scott Reed, Jack Rae, Chris Dyer, and Phil Blunsom. Neural arithmetic logic units, 2018.
[126]Trieu H. Trinh and Quoc V. Le. A simple method for commonsense reasoning, 2019.
[127]John von Neumann. *Mathematical Foundations of Quantum Mechanics*. Springer, 1932.
[128]Mei Wang, Xin-You Lu, Ying-Dan Wang, J. Q. You, and Y. Wu. Macroscopic quantum entanglement in modulated optomechanics, 2017.
[129]Rui Wang, Joel Lehman, Jeff Clune, and K. Stanley. Paired open-ended trailblazer (poet): Endlessly generating increasingly complex and diverse learning environments and their solutions, 2019.
[130]J. A. Wheeler. *Geons, Black Holes, and Quantum Foam: A Life in Physics*. W.W. Norton & Co, 1998.
[131]Frank Wilcz. Quantum time crystals, 2012.
[132]Frank Wilczek. Physics in 100 years, 2015.
[133]Frank Wilczek. Theory vision. In *LHCP 2016*, 2016.
[134]A. Wiles. Modular elliptic curves and fermat's last theorem. *Ann. of Math.*, 141(2–3):443–551, 1995.
[135]Jiasong Wu, Ling Xu, Youyong Kong, Lotfi Senhadji, and Huazhong Shu. Deep octonion networks, 2019.
[136]Xian Wu, Baoxu Shi, Yuxiao Dong, Chao Huang, and Nitesh V. Chawla. Neural tensor factorization. In *WSDM'19 (Proc. of the 12th ACM International Conference on Web Search and Data Mining)*, 2019.
[137]Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov, and Quoc V. Le. Xlnet: Generalized autoregressive pretraining for language understanding, 2020.
[138]T. Yoneya. Wandering in the fields. *Physics Review Letter*, 78(1219), 1997.
[139]Xiangyu Zhang, Xinyu Zhou, Mengxiao Lin, and Jian Sun. Shufflenet: An extremely efficient convolutional neural network for mobile devices, 2017.
[140]M Zheng, A Allard, P Hagmann, and M Á Serrano. Geometric renormalization unravels self-similarity of the multiscale human connectome. 2019.
[141]Pan Zhou, Xiao-Tong Yuan, and Jiashi Feng. Faster first-order methods for stockastic non-convex optimization on riemannian manifolds, 2018.
[142]Tianyi Zhou, Dacheng Tao, and Xindong Wu. Manifold elastic net: a unified framework for sparse dimension reduction. *Data Mining and Knowledge Discovery*, 22:340–371, 2010.
[143]Barret Zoph and Quoc V. Le. Neural architecture search with reinforcement learning, 2016.
[144]W. H. Zurek. Quantum darwinism and envariance, 2003.