# Implicit Multi-Fold Mechanisms in a Neural Network Model of the Universe

Stephane H. Maes[1]

September 12, 2020

**Abstract:**

*In a multi-fold universe, gravity emerges from Entanglement through the multi-fold mechanisms. As a result, gravity-like effects appear in between entangled particles that they be real or virtual. Long range, massless gravity results from entanglement of massless virtual particles. Entanglement of massive virtual particles leads to massive gravity contributions at very smalls scales. Multi-folds mechanisms also result into a spacetime that is discrete, with a random walk fractal structure and non-commutative geometry that is Lorentz invariant and where spacetime nodes and particles can be modeled with microscopic black holes. All these recover General relativity at large scales and semi-classical model remain valid till smaller scale than usually expected. Gravity can therefore be added to the Standard Model. This can contribute to resolving several open issues with the Standard Model without new Physics other than gravity. These considerations hints at a even stronger relationship between gravity and the Standard Model.*

*Recently a controversial series of papers ended up proposing the possibility that the universe be a neural network. It is the result of observing that with an irreversible thermodynamics model of the learning process of the neural network, it might appear possible to model quantum and classical physics, to observe the emergence of a General Relativistic spacetime with gravity, and plausible to construct a generalized holographic principle beyond the AdS/CFT correspondence conjecture. The approach has been received with some skepticism.*

*In this paper, we do not try to assess the validity of the approach and proposal. We simply assume that the proposal amounts to showing that neural network (NN) learning with a suitable thermodynamically related loss function (aka cost function) optimization, that amounts to extremize the free energy of the system, can model the Physics of the universe. When we add a model of entanglement, we discover that the neural network must allow its involved neurons to pair into pairs (or groups) of (dynamic) Qubits. Non quantum NN neurons cannot be simply grouped this way. Instead one need to add new (external) NN, that themselves emulate Qubit behaviors, between the "entangled" nodes. It amounts to match the multi-folds, including their spacetime extensions, and mechanisms. Furthermore the additional NN, explain the possibility to induce 7D physics in 4D space time to induce the Standard Model with gravity ($SM_G$), encountered with multi-fold universes, while the multi-fold dynamics itself (in AdS(5), does not have necessarily have to be governed by General Relativity.*

*The work also leads us to wonder if Quantum Physics is fundamental or emergent; especially with what we already know about entanglement and spacetime construction by random walks, in multi-fold universes.*

*An appendix also discusses how NN models could relate to the Wigner's wonder at why mathematics describe the Physical world.*

____

## 1. Introduction

[1] shmaes.physics@gmail.com

Cite as: Stephane H Maes, (2020), "Implicit Multi-Fold Mechanisms in a Neural Network Model of the Universe", viXra:2012.0191v1, https://vixra.org/pdf/2012.0191v1.pdf, https://shmaesphysics.wordpress.com/2020/09/12/implicit-multi-fold-mechanisms-in-a-neural-network-model-of-the-universe/, September 12, 2020.

The new preprint [1] proposes contributions to several open problems in physics like the reconciliation of General Relativity (GR) with Quantum Physics, explaining the origin of gravity proposed as emerging from quantum (EPR-Einstein Podolsky Rosen) entanglement between particles, detailing contributions to dark matter and dark energy and explaining other Standard Model mysteries without requiring New Physics beyond the Standard Model other than the addition of gravity to the Standard Model Lagrangian. All this is achieved in a multi-fold universe that may well model our real universe, which remains to be validated.

With the proposed model of [1], spacetime and Physics are modeled from Planck scales to quantum and macroscopic scales and semi classical approaches appear valid till very small scales. In [1], it is argued that spacetime is discrete, with a random walk-based fractal structure, fractional and noncommutative at, and above Planck scales (with a 2-D behavior and Lorentz invariance preserved by random walks till the early moments of the universe). Spacetime results from past random walks of particles. Spacetime locations and particles can be modeled as microscopic black holes (Schwarzschild for photons and spacetime coordinates, and metrics between Reisner Nordstrom [2] and Kerr Newman [3] for massive and possibly charged particles – the latter being possibly extremal). Although surprising, [1] recovers results consistent with other like [4], while also being able to justify the initial assumptions of black holes from the gravity or entanglement model in a multi-fold universe. The resulting gravity model recovers General Relativity at larger scale, as a 4-D process, with massless gravity, but also with massive gravity components at very small scale that make gravity significant at these scales. Semi-classical models also turn out to work well till way smaller scales that usually expected.

The present paper discusses how modeling entanglement by extending the neural networks (NN) proposed to model Physics in the universe as neural networks theories [5,6,7], can be seen as recovering key results of [1] and [9].

In this paper, we remain at a high level of discussion of the analysis and references are generic for the subjects. It makes the points accessible to a wider audience and keeps the door open to further papers or discussions devoted to details of interest. Yet, it requires the reader to review [1], as we do not revisit here all the details of the multi-fold mechanisms or reconstruction of spacetime. More targeted references for all the material discussed here are compiled in [1] and derived papers.

## 2. NN Model of the Universe

[6] shows that if information theory is modeled with (covariant) irreversible / non-equilibrium thermodynamic processes then, close to equilibrium, the conjugate thermodynamics variables of the information content (tensor) is an emerging spacetime following the Hilbert Einstein spacetime. This result is to be related to [8], that derives emergence of quantum mechanics from classical irreversible thermodynamics. Away from equilibrium, the picture is less clear. We note that the irreversibility has to be directly related to the quantum behavior.

Following up on these results, [7] proposes a thermodynamics model for Machine Learning (ML) and derives a proposal for Thermodynamics of learning. NN are example of ML, but we know that any AI or ML algorithm can always be modeled as a NN [17,18,22].

[5] then models NN thermodynamics, using [7] and inspired from [6,8] and shows:

- Close to equilibrium and when the entropy contributions from learning are small, one can recover a Schrödinger equation and a wave function that results from the stochastic dynamics of the training variables randomly trying to find where to go to learn. It amounts to small scale events, trying different evolution to find hints of the best ones, not really changing much with respect to what the NN has learned, and it denotes a state of the NN, where equilibrium has been reached, and new variables values for the models are randomly visited just in case they could help or because learning continues.

- Further away from equilibrium, where random fluctuations of the $q_i$, the learning variables, are smaller and less visible, and hence at larger scales, and therefore when learning process dominates the thermodynamic, the training variable have an evolution that can be characterized by a classical Hamiltonian and therefore can be modeled by classical Physics. It corresponds to a state of the NN, where it can estimate how to progress to learn or improve the loss/cost function (think of gradient like steepest gradient descent methods for learning/training/optimization).
- When modeling directly the dynamics of the state of the neurons, [6] applies and under suitable conditions (close to equilibrium and with weak interaction between the neurons (at least when nonlocal)), the dynamics of the neurons follows Einstein's GR field equations.
- Analyzing In and Out layers of the NN versus hidden layers, one can hypothesize ways to recover a generalized holographic principle that would link a quantum mechanically dominated NN (In + Out layers) to a deep / many layers NN dominated by gravity.

Appendix A presents additional considerations on what we can learn from [5].

However, this model does not model entanglement yet. It is a key missing part before we can claim to have a truly complete quantum model emerging from [5].

# 3. Adding Entanglement to NN

Qubit/Quantum NN and Fuzzy Logic are traditional ways to add quantum entanglement effects [10,11,12]. Essentially the most natural way is to allow neurons to now be Qubits, instead of, say binary neurons.

But this is not what we are discussing here. The NN in [5] are not modeled as Qubits. Some additional considerations should probably be added to [5] to handle these, although may be without significant impact. However that is different from having nodes switching from neurons to Qubit based neurons. That is not trivially handled by the Quantum NN in [5]. In future work, we will show it actually could be handled by a different NN, as [17,18,22] suggest that a different NN could do the job, unfortunately possible at the cost of an excessive cost in complexity and number of neurons and variables. This option (*) is not discussed in the present paper and will be the topic of a future paper.

Returning to [5], entanglement results from having regions of the wave function entangled, corresponding say to different EPR particles. These correspond to different hidden values $x_i^{(K)}$ where K denotes regions not directly interacting (and as such non-local).

Per [5] we know that, before entanglement, the $x_j^{(L)}$ form a spacetime governed by GR (at least when interactions between the neurons is weak). Entanglement (and disentanglement) are strongly interacting disrupting events that correlate and bind the behaviors between two (or more) different $x_i^{(O)}$ and $x_j^{(P)}$, or disrupt such correlation. Entangled region become essentially a self-interacting system with, at best, weak (as in with small interaction impacts) interactions with the rest. Changes to the training variables are correlated so that they impact the hidden variable consistently with the entanglement behavior. Such an entanglement amounts to having $x_i^{(O)}$ and $x_j^{(P)}$ now behaving like a Qubit. To be handled with NN à la [5], we need an external NN grafted between the now entangled hidden variable to emulate a Qubit. Such a NN is for example studied in [13]. It is possible to show that any NN with such grafts is in fact a QNN or hybrid NN + QNN. A more detailed discussion and proof will be provided in a future dedicated paper. Of course, other approaches may exist.

As a result, when entanglement takes place, a discontinuity in the wave function/equations is reflected by the grafts of additional NNs. This NN also follows [5] and so an extra spacetime appears between the two entangled hidden variable regions (the rest remaining the same). The grafting process is most probably not governed by GR, but once connected, GR applies on the resulting spacetime + extra spacetime, which modifies the gravity felt, per

[5], on the pre-existing spacetime (not from the grafted NN). At disentanglement, we revert to a previous unentangled topology, and we can consider that $x_i^{(O)}$ and $x_j^{(P)}$ also maintain their place in the main spacetime. Of course, all these steps have impact on the entropy production and destruction and the free energy.

As time passes, the grafted NNs evolve to remain connected between the entangled points and new NNs can be grafted, seeded by previous values. Older ones remain in place. When disentanglement takes place, they are removed.

# 4. Multi-folds and Multi-fold Mechanism in Entanglement of Quantum systems described by NN

Fundamentally, section 3 depicts the multi-fold mechanisms proposed and detailed in [1]: extra folds are made available for path integral paths of the entangled particle and result into gravity impact on the background spacetime (as attractive gravity like effective potentials or effective curvature between entangled systems [14]). We recover the extra gravity due to the folds, here exemplified by the interactions in the grafted spacetime that curves as a result. Mappings effects result from older grafted NN remaining in place till disentanglement, located around/between the neurons and affected by the correlation between entangled neurons. As the graft is done via another process, it is not expected to follow GR, as we suggested for the dynamics of multi-folds [1,16].

We also recover the additive effect of multiple sources in multi-fold: if different source merge the different additional grafted NNs add their effects.

The disentanglement process, where we remove the graft, is related in our view to how we handle fold deactivation to maintain unitarity as well as the process of deactivation.

The irreversibility associated to Quantum Physics [8] is also interesting considering that [1] predicts that multi-fold mechanisms are a source of irreversibility (deactivation can't "grow" as activation starting from local points per the hierarchical principle) and T-symmetry violation. We expect to explain it beyond just disentanglement in future work. Only allowing entanglement to grow from neighboring points, is a way to enforce the hierarchical principle discussed in [1].

# 5. Recovering the Standard Model with Gravity ($SM_G$)

[9] did all the work to recover from [1] the Standard Model with gravity: $SM_G$. To do so, it relied on induced space-time-matter from locally embedding spacetime into a 7D unconstrained (i.e. non compact) Kaluza-Klein Universe.

The result here show how "entangled" points of spacetime are locally embedded is a bigger spacetime (with 2 times the space dimensions, i.e. 7D for a 4D background spacetime), therefore also relating entangled NN with the proposal of [9] for induction from 7D Physics.

# 6. AdS(5) in the NN model of the Universe

Because additional NN are grafted, not generated by optimization of just the original NN (at least within the context of (*)), their dynamics and kinematics are not a priori governed by GR (not forbidden, as easily seen when considering approaches beyond (*)), but not implied). We recover our result from (and a priori a difference with ER= EPR) [9,24].

Repeating our multi-fold analysis that describes the quantum origin gravity, from entanglement via the multi-fold mechanisms, results into an AdS(5) dual space surrounding every spacetime point and generated by the folds [1]. That is for now beyond the model of [5], yet fully compatible.

The proposal of [5] in terms of Holographic duality can not only model and extend the AdS/CFT correspondence conjecture but also the factual correspondence discussed in [1,9,15,16]. Maybe the extensions proposed in [5], for going beyond the zone of applicability of Quantum Physics, also apply to multi-fold universes. It could be worth further investigation.

# 7. Quantum Physics, Irreversibility and Equilibrium

The recovery of GR close to equilibrium, in [5], agrees with the results from many other works, following the pioneering recovery of GR, by Ted Jacobson, who applied Thermodynamics to spacetime treated as adiabatic, and in, or close to, equilibrium. That work was used in later work to study quantum gravity and entanglement entropy. Indeed this adiabatic and equilibrium regime is the domain of applicability of GR and of Quantum Physics

In general, the equations of conventional Quantum Physics appear time reversible. [1] showed that gravity and entanglement is not T-symmetric. It hints at why [8] can model quantum physics and phenomena by thermodynamics of irreversible systems. And irreversibility is expected to result from purely quantum effects, What is more representative of quantum Physics than entanglement and disentanglement?

But [8] provides another result that is possibly even more important: if thermodynamics is such a good model, then Quantum Physics may rather be an emergent theory and we still need to find the more fundamental underlying theory. As [1] provide both idea of irreversibility of entanglement and construction of spacetime via random walks, these may be good starting point. It will be object of future work.

*12/24/20 note: See [23] for such a follow-up contribution where the W-type hypothesis introduced in [23] complements [1] to provide glimpses of such a fundamental theory. Interestingly it also extends causes for T-asymmetry to wave function collapses.*

# 8. Conclusions

There have been already many hints of relationships between spacetime, entanglement, thermodynamics and information theory like treating the universe as universal Quantum Computer, encountering error correcting code in spacetime (including in [1]), deriving GR from spacetime properties in equilibrium and the relationships between gravity, entropy and entanglement entropy as well as the principle of conservation of information in Quantum Physics and the information paradox with Black holes. Information and Physics are closely related and this paper, along with many of its references, add to these observations.

In this paper, we did not try or claim to validate or endorse the proposal that the universe would be a NN. We rather started from the point of view that Physics and the Universe seems to follow models analog to the evolution of a learning NN using the models of [5]. Adding entanglement to the models, we immediately recover strong hints of the multi-fold mechanisms: a way to add it as a model can be interpreted immediately as adding multi-folds with all the implications of [1,17]; including in particular compatibility with the approach that allowed us to recover the $SM_G$, induced from a 7D unconstrained KK universe. It certainly reinforces the claim that multi-fold universes should be seriously considered.

On the flip side, we showed compatibility of multi-fold universes with the proposal of [5], extended with NN grafts that model entanglement, something that [5] must add to its model to further its claims about the universe by supporting entanglement.

We also encountered hints that Quantum Physics may be an emergent theory, if so well modeled by Thermodynamics. Being a NN maybe an explanation. Other explanations, more physical, may exist.

# Appendix A – A NN model of the world? An alternate interpretation

*(This appendix will also be repurposed into a dedicated paper not tied to multi-fold universes)*

[5] proposes that the universe is a NN. We do not believe that this is the only interpretation of the results presented in [5] and we want to propose an alternative explanation. As already mentioned, the NN approach can be seen as a model of the dynamics of Physics in the universe. Such model is mathematical, in fact it is a consequence of Hilbert 13th problem and the ability to model any system with deep hidden layers and in particular NN as demonstrated with the Kolmogorov-Arnold representation Theorem [17] and the Universal approximation theorem [18].

In the present case the dynamics of the state variables, i.e. the equation of motion, are the approximated functions. Per the theorems above we know such approximation is (almost) always possible (up to discontinuities) and to any desired degree of accuracy (for the right optimization strategy in the case of NN).

What is interesting, is that if the algorithm for loss/cost function optimization relies on (classical) Thermodynamics (for Irreversible and for non-equilibrium processes with a Free Energy model), it uncovers naturally the dynamics described in section 2 [5], where the fact that the NN includes also the model of the learning processes allows to capture in one shot dynamics of the physical system (i.e. the universe) and the dynamic of information processing; therefore concretizing the physical information theory aspects also (e.g. see [19] for related aspects of physical information theory); something that now can be captured into a common Thermodynamics (and physical) model. It goes beyond [8] and justifies considerations like Learning's Thermodynamics or the principle of conservation of information. In our view, much more than having a NN modeling (or being per [5]) the universe, the key aspect is that we have a complete model for physical and information entropy modeling and computing.

In such a model, it makes sense that entropy extremization and action extremization become equivalent or dual. It is also natural to see that, at small scales, quantum fluctuations around equilibrium imply fluctuations of the learning variables, and the NN state, while at larger scales away from equilibrium (albeit still close), the system will rather behave classically as a learning system (to go back to equilibrium).

So we interpret [5] as a model that shows first and foremost how Physics + Information Theory coexist into a larger model. The model of [5] has its own dynamics. These dynamics may be seen as a model of how physical systems like the universe handle information conservation or just as an algorithm to derive the same outcome. More work is needed to determine that. If it is the former, this may actually be a way to answer why and how mathematics are so good at modeling the Universe as asked famously by Wigner [20], and others, and it would be aligned with Tegmark's view [21]. Indeed, [5] would now amount to modeling how the universe remains close to thermodynamic equilibrium while always reacting to changes and fluctuation (e.g. random, thermal external, etc.) to catch up with the mathematical prescription aiming at optimizing the loss/cost function while evolving with minimum disruptions as captured by extremization of the entropy and action changes: physical systems take some "guessed optimized efforts" to catch up and follow the mathematics that describe them correctly and these mathematics are the reflection of this process. It is a direct application of Pontryagin's maximum principles and theorem [25-27].

____

**References:**

[1]: Stephane H. Maes, (2020), "Quantum Gravity Emergence from Entanglement in a Multi-Fold Universe", viXra:2006.0088v1, https://vixra.org/pdf/2006.0088v1.pdf (June 9, 2020).
[2]: https://en.wikipedia.org/wiki/Reissner%E2%80%93Nordstr%C3%B6m_metric
[3]: https://en.wikipedia.org/wiki/Kerr-Newman_metric
[4]: Burinskii, Alexander, (2008), "The Dirac-Kerr-Newman electron", arXiv:0507109v4
[5]: Vitaly Vanchurin, (2018), " The world as a neural network", arXiv:2008.01540v1
[6]: Vitaly Vanchurin, (2018), "Covariant Information Theory and Emergent Gravity", arXiv:1707.05004v4
[7]: Vitaly Vanchurin, (2020), "Towards a theory of machine learning", arXiv:2004.09280v3
[8]: D. Acosta, P. Fernandez de Cordoba, J. M. Isidro, J. L. G. Santander, (2012), "Emergent quantum mechanics as a classical, irreversible thermodynamics", arXiv:1206.4941v2
[9]: Stephane H Maes, (2020), "Tracking Down The Standard Model With Gravity In Multi-Fold Universes", viXra:2011.0208v1, https://shmaesphysics.wordpress.com/2020/08/30/tracking-down-the-standard-model-with-gravity-in-multi-fold-universes/, August 20, 2020.
[10]: https://en.wikipedia.org/wiki/Quantum_neural_network
[11]: Nobuyuki Matsui, Haruhiko Nishimura and Teijiro Isokawa, (2009), "Qubit Neural Network: Its Performance and Applications", in Tohru Nitta, (2020), "Complex-valued Neural Networks: Utilizing High-dimensional Parameters", Information Science Reference
[12]: Gopathy Purushothaman and Nicolaos B. Karayiannis, (1997), "Quantum Neural Networks (QNN's): Inherently Fuzzy Feedforward Neural Networks", IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 8, NO. 3, MAY 1997
[13]: Emmanuel Flurin, Leigh S. Martin, Shay Hacohen-Gourgy, Irfan Siddiqi, (2020), "Using a Recurrent Neural Network to Reconstruct Quantum Dynamics of a Superconducting Qubit from Physical Observations", PHYSICAL REVIEW X 10, 011006
[14]: Stephane H Maes, (2020), “Gravity-like Attractions and Fluctuations between Entangled Systems?”, viXra:2010.0010v1, https://shmaesphysics.wordpress.com/2020/06/25/gravity-like-attractions-and-fluctuations-between-entangled-systems/, June 24, 2020.
[15]: Stephane H Maes, (2020), “Area Laws Between Multi-Fold Universes and AdS”, viXra:2010.0207v1, https://shmaesphysics.wordpress.com/2020/08/10/area-laws-between-multi-fold-universes-and-ads/, August 10, 2020.
[16]: Stephane H Maes, (2020), “Multi-fold Gravitons In-N-Out Spacetime”, viXra:2010.0155v1, https://shmaesphysics.wordpress.com/2020/07/27/multi-fold-gravitons-in-n-out-spacetime/, July 27, 2020.
[17]: Wikipedia, "Kolmogorov–Arnold representation theorem" https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Arnold_representation_theorem, Retrieved on September 14, 2020.
[18]: Wikipedia, "Universal approximation theorem", https://en.wikipedia.org/wiki/Universal_approximation_theorem, Retrieved on September 14, 2020.
[19]: Seth Lloyd, (2006), "Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos", Alfred A. Knopf
[20]: Wigner, E. P. (1960). "The unreasonable effectiveness of mathematics in the natural sciences. Richard Courant lecture in mathematical sciences delivered at New York University, May 11, 1959". Communications on Pure and Applied Mathematics. 13: 1–14.
[21]: Max Tegmark, (2007), "The Mathematical Universe", arXiv:0704.0646v2
*(Added when pre-print was published on vixra.org)*
[22]: Andre Ye, (2020), "Every Machine Learning Algorithm Can Be Represented as a Neural Network", https://towardsdatascience.com/every-machine-learning-algorithm-can-be-represented-as-a-neural-network-82dcdfb627e3. Retrieved on December 19, 2020
[23]: Stephane H Maes, (2020), “The W-type Multi-Fold Hypothesis and Quantum Physics Interpretation of wave Functions and QFT”, viXra:2207.0118v1, https://shmaesphysics.wordpress.com/2020/12/24/the-w-type-multi-fold-hypothesis-and-quantum-physics-interpretation-of-wave-functions-and-qft/, December 20, 2020.
[24]: Maldacena, Juan and Susskind, Leonard (2013). “Cool horizons for entangled black holes”. Fortsch. Phys. 61 (9): 781–811. arXiv:1306.0533

[25]: Wikipedia, "Pontryagin's maximum principle", https://en.wikipedia.org/wiki/Pontryagin%27s_maximum_principle. Retrieved on September 29, 2020.
[26]: "13 Pontryagin’s Maximum Principle", http://www.statslab.cam.ac.uk/~rrw1/oc/L13.pdf. Retrieved on September 29, 2020.
[27]: Thayer Watkins, "The Nature of the Principle of Least Action in Mechanics", https://www.sjsu.edu/faculty/watkins/minprin.htm. Retrieved on September 29, 2020.