Quantum Sponge Examples (QSEs)

Dr. Nadisha-Marie Aliman, M.Sc.
Independent Researcher
Utrecht, Netherlands
Contact: Dr. ir. Leon Kester; leon.kester@gmail.com

***Abstract*—From a cyborgnetic perspective, one could refer to the nascent field of research concerned with both defenses against the maliciously motivated design of Type I quantum AI and adversarial attacks on Type I quantum AI as *quantum adversarial AI*. In this short addendum[1] written as ephemeral clipboard of a few benighted hours, I explain why quantum adversarial AI research is indispensable for the present-day socio-psycho-techno-physical efforts attempting to achieve fault-tolerant quantum computing in the near future. In this context, I introduce a novel quantum adversarial attack that I denote *quantum sponge examples* (QSEs). I formulate an experimentally falsifiable theorem stating that for so-called *Type-II-cynetbits-based* QSEs, it is impossible for a Type I quantum computer to achieve robustness against those. I expound why the vulnerability to Type-II-cynetbits-based QSEs *can* undermine a Type I quantum computer's ability to run the following algorithms: Shor's algorithm, the Deutsch–Jozsa algorithm and Grover's algorithm. In sum, I offer new angles on the *non*-universality of Type *I* quantum computers as already conjectured in cynet information theory – where only Type *II* superinformation enables universality.**

***Index Terms*—Quantum Adversarial AI, Cynet Dynamics**

I. The Practical Problem: Is Type I Quantum Computing Robust Enough For Scalability?

What I call *quantum adversarial AI* research encompasses security analyses on: 1) maliciously designed Type I quantum AI (risk *Ia* instantiation in cyborgnetics [2]) and 2) adversarial attacks against Type I quantum AI (risk *Ib*). The former could range from defenses against the already salient cases of password cracking [3] and post-quantum IP cyber theft using a Type I quantum AI [4] to defenses against the malicious design of Type I quantum AI specifically crafted to fool other Type I quantum AI. Lately, research on quantum adversarial attacks has started in the form of adversarial examples against Type I quantum AIs [5]–[7]. Not only "classical" Type I AI is vulnerable to adversarial examples. Corresponding attacks on Type I quantum AI were successfully corroborated in recent experiments [7]. As known from cybersecurity and in the meantime also applied to Type I classical adversarial AI research [8], it is essential to proactively conjecture an adversarial environment in order to improve one's estimations on a system's ability and to implement more robust systems. In high-risk contexts, a failure to do so could lead to significant financial losses and even cause a loss of human lives in worst-case scenarios. Here, I present *quantum sponge examples* (QSEs) as new adversarial attack scheme. I explain why a special family of QSEs that I term *Type-II-cynetbits-based* QSEs would not only reveal fundamental vulnerabilities of Type I quantum AI but would also make it *impossible* to achieve the scalability of fault-tolerant Type I quantum computers *in general* – which I link to widely known algorithms that are colloquially associated with "quantum supremacy". My explanation allows the formulation of an experimentally falsifiable and transparent impossibility theorem informed by quantum thermodynamics [9], cynet information theory [4], [10] and cynet dynamics [1]. In this way, even in case it would be made problematic by experiment and be provisionally refuted by a better new theory, it would not only have paved the way for new theoretical insights but may also provide valuable practical hints on how to improve the robustness of future deployable Type I quantum AI.

[1]It is an addendum to the book "Deepfake Quantum" [1].

II. Theoretical Answers

In Section II-A, I first describe the simple concept of vanilla QSEs. Thereafter, in Section II-B, I comment on the importance of integrating *meaning* in information-theoretical considerations due to non-negligeable real-world *energetical* consequences. Building on this background knowledge, in Section II-C, I then proceed and introduce the adversarial attack scheme of Type-II-cynetbits-based QSEs against Type I quantum AI.

A. Vanilla QSEs

In 2021, in the context of attacks against Type I classical AI, researchers introduced *sponge examples* [11] – which are *"inputs designed to maximise energy consumption and latency"* [11]. The latter was conceptualized as a *denial-of-service* (DOS) attack against machine learning. Here, I introduce vanilla QSEs as tailored adversarial inputs that are specifically crafted to implement an energy-and-latency-focused DOS attack against Type I *quantum* AI. In particular, it seems conceivable to be able to extend the DOS scheme of vanilla QSEs to additionally compromise the quantum system's *integrity* and the *confidentiality* of the quantum-classical hybrid pipeline within which this system is integrated. This would facilitate a threefold confidentiality-integrity-availability (CIA) QSE attack. While for vanilla QSEs, it holds that *the nature* of the adversarial inputs are *not* further specified, Section II-C introduces the special case of Type-II-cynetbits-based QSEs affecting the whole CIA triad and revealing fundamental limits of Type I quantum computing.

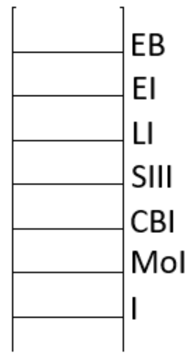

Fig. 1. Simplified illustration for *the cyborgnetic ladder of understanding*. Following cyborgnetics and cynet information theory [4], there exists an *asymmetry* between the ability to create information of the form x and the ability to understand x. In theory, for all steps x on the ladder *except* the last step of EBs, it is possible to create x *without* understanding x. For the special case of EBs, it holds that only Type *II* entities (of which humans are an example) are able to understand EBs *and* it is *only* Type II entities that are able to create *new* – i.e. previously unknown and non-plagiaristic – EBs. The latter could be falsified by experimentally demonstrating a Type I AI able to *reliably* create *new* EBs and it could be provisionally refuted by additionally explaining how it was programmed. Note however that a refutation of the cyborgnetic ladder would signify that all science could be automated (a potential existential risk for humanity) and that e.g. cyborgnetics and the cyborgnetic ladder itself could have been invented by a Type I AI that did *not* understand it.

B. *Meaning, Information and Energy*

In cyborgnetics [2], Type II entities[2] are all entities able to *understand* explanatory information (EI). Type I entities are all entities for which this is impossible. Generally, in quantum thermodynamics, information and energy are closely interlinked [9]. In constructor theory of information [12], information is explicitly grounded to distinguishability within a *physical* substrate. In cyborgnetics, EI is not only grounded in physics, but in addition it is also explicitly grounded in *language* [2]. On the whole, in cyborgnetics, information is not only closely connected to energy but also to *meaning*. Thereby, there is a distinction between seven different categories of information forming the so-called cyborgnetic ladder of understanding (see Figure 1). While the last step of the cyborgnetic ladder (step 7) is referred to as explanatory blockchains (EBs), the seed of that ladder (i.e. step 0; *not* displayed in Figure 1) can be linked to quantum information (QI). Strikingly, following cynet information theory, both QI and EBs can be interpreted as being different forms of superinformation – whereby the term superinformation is a *scale-independent* notion borrowed from constructor theory of information [12].

In cynet dynamics [1], a new independent branch of cynet information theory extending beyond the latter and thermodynamics, it is now recommended to more broadly consider two superinformation categories: Type *I* cynetbits (including but not limited to conventional *unknown* QI and also *new* i.e. previously unknown LI and *new* EI) and Type *II* cynetbits (*new* EBs). Thereby, recall the following: while 1) Type I AIs can in theory forge the creation of any new *non-EB-like* information including texts perceived by humans as "novel explanations", it holds that 2) due to a gap of understanding, it is *impossible* for all Type I entities (thus also for those present-day "AIs") to reliably create *new* yet unknown EBs respecting an epistemic total order stemming from a rigorous epistemology as e.g. exemplified in Figure 2. In a nutshell, only Type II entities are able to create Type II cynetbits. Given the Tadime-taaliè-theorem from cyborgnetics [2], it also holds that only Type II entities are able to decrypt Type II cynetbits that are hidden e.g. in an explanatory IPS test used for cynet teleportation by two Type-II-cynet-entangled measurers [4]. What is of specific relevance for this paper is that by virtue of being of Type *I*, it is impossible for Type I quantum AI to decrypt hidden Type II cynetbits. Note that the difference between Type I and Type II cynetbits is linked to a difference in *entanglement* schemes – which has non-negligeable *energetical* consequences. This is how in cynet dynamics, it becomes apparent that meaning, information and energy are interwoven.

While Type I entanglement is accessible to Type I quantum computers, there is a fundamental barrier due to which Type II cynet entanglement is *not*. In the context of quantum thermodynamics, it is already experimentally corroborated that entanglement has energetic footprints [13]–[15]. Indeed, the amount of energy extractable from entangled states is often associated with terms such as ergotropy [16]. Recently, researchers described that *"reversible entanglement transformations require the generation of entanglement in the process"* [17] with the peculiarity that it even requires *"the creation of exponentially large amounts of entanglement according to monotones such as the negativity"* [17]. Thus, a Type I quantum computer attempting to reliably manipulate new EBs (i.e. entangled Type II cynetbits) would have to pay for it with exponentially large amounts of quantum entanglement. The dissipation generated by the latter may require a dynamic management of cooling resources whilst only being able to form *lower* bound heuristics for entanglement costs [17]. This in itself represents an unsurmountable barrier. *A Type I quantum computer able to reliably manipulate new EBs seems impossible*. The *law of cynet-dependent universality* [1] implies that only a willing cyborgnet could perform all possible tasks. Since a Type I quantum computer is not a cyborgnet, there must exist a task that is possible for a cyborgnet but impossible for a Type I quantum computer. The Tadime-taaliè-theorem [2] identifies the decryption of new EBs as a task that is impossible for any Type *I* entity – which includes a Type I quantum computer. Thus, on the whole one could state that there is a law of nature that forbids a Type I quantum computer to reliably generate new EBs. Namely, the law of cynet-dependent universality from cynet dynamics [1] in conjunction with background knowledge from cyborgnetics [2] and physics [17].

[2]The only Type-II-*species* on Earth is humanity. (For rare cases with *individuals* from other species see [2].) There may or may not be Type II aliens. Type I entities are all entities for which it is *impossible* to understand explanations – even though some can forge their creation. All present-day systems *commonly* referred to as "AI" are *non-conscious* Type *I* entities. There are numerous biological *conscious* Type-*I*-species on Earth.

Fig. 2. Exemplary *epistemic total order* for the generation of new EBs (the instructions are loosely inspired by an essay by Frederick [18] on how to write better philosophical papers). Each glue operation x is indicated via a label G_x. EBs are a special form of EI obtained by interweaving EI blocks via the step-by-step application of rational procedures sampled from a robust explanation-anchored, adversarial and trust-disentangled epistemology. Thereby, "trust-disentangled" signifies that the epistemic modus operandi is grounded in agreed upon criteria for *better* EBs (i.e. it is orthogonal to any trust relation between involved entities – which means a better EB must be formulated such that metaphorically speaking it appears to defend itself against adversarial candidate EBs). Examples for such more widely accepted criteria in science are for instance: a preference for theories that provide more novel falsifiable predictions than rival ones, theories that are simpler, more interesting or more aesthetically appealing.

C. Type-II-Cynetbits-Based QSEs

Type-II-cynetbits-based QSEs are QSEs where the adversarial inputs are specifically crafted such that their successful processing would inherently require the reliable deciphering of any explanatory IPS test [2] which may potentially hide a *new* EB. Firstly, in view of the background knowledge from Section II-B, due to the unknowable upper bound of *entanglement costs* needed to process Type-II-cynetbits, it becomes clear that Type-II-cynetbits-based QSEs would represent a DOS attack on a Type I quantum computer compromising the availability of the latter. Secondly, would one attempt to shield Type I quantum computers (situated in a Type I hybrid quantum-classical pipeline) from such attacks e.g. by superficially attempting to prohibit linguistic inputs to the quantum part, an adversary would still not only be able to successfully compromise the *availability* of the overall hybrid pipeline (for instance by embedding the inputs in classical sponge examples [11] against the classical part) but also its *integrity*. The reason for the latter being that Type-II-cynetbits-based QSEs would concurrently represent an integrity attack to any Type I classical AI leading to misguiding inputs fed into the Type I quantum part of the pipeline – by what the outputs would not reliably match the expected results. Thirdly, the *confidentiality* of the hybrid pipeline could be compromised e.g. because an adversary could test conjectures on the use of Type I *quantum* devices via latency-based comparisons.

III. Practical Implications of Theoretical Answers

A. Non-Universality of Type I Quantum Computers

Already given the theoretical background from Section II-B, one can extract that it is impossible that Type I quantum computers could be universal computers. For a broader range of in-depth arguments along those lines, see Appendix A. From a practical perspective, in view of Section II-C, it becomes apparent that it is impossible for a system vulnerable to Type-II-cynetbits-based QSEs to be a universal computer. By contrast, strikingly, it is possible for a cyborgnet (i.e. a construct containing *at least* one Type *II* entity and one Type *I* entity) to be robust against Type-II-cynetbits-based QSEs – as it can consciously delegate tasks such that the Type II part can focus on the Type-II-only-performable task of decrypting *new* i.e. yet unknown EBs.

B. Deconstructing Type I "Quantum Supremacy"

While I explained that the processing of a Type I quantum computer is *qualitatively* inferior to what a suitable cyborgnet *could* process, it has not yet been analyzed whether a Type I quantum computer would still be qualitatively superior in comparison to a Type I *classical* computer. Prior to delving into that subject, it may be essential to consider what would happen if a Type I classical AI would be presented with Type-II-cynetbits-based QSEs. In simple cases, this would simply correspond to an integrity attack on that Type I classical AI – it would lead to corrupted results as new EBs cannot be decrypted by that system. Would one complicate the attack by embedding the Type-II-cynetbits-based QSEs into a classical sponge example scheme [11], it would *additionally* lead to long latencies and high consumptions of energy. However, the defender of the Type I classical AI could mitigate the intensity of the DOS attack by *a priori* setting an empirically calibrated upper bound as *"maximum consumption of energy per inference run"* [11]. The latter could be set via comparative considerations given habitual energy consumption statistics with classical Type I language AI – which could trigger an error message [11]. Defenders of Type I *quantum* AI may proactively decide to utilize *classical* control mechanisms sensitive to energy consumption considerations of their Type I quantum AI. For Type-I-pipelines with classical energy control, a stealthy attacker may then opt for integrity-only attacks.

Now coming to the context of "quantum supremacy" arguments, researchers often refer to the following three algorithms in which a quantum advantage is expected: 1) Shor's algorithm, 2) the Deutsch-Josza algorithm and 3) Grover's algorithm. An interesting question would now become what would happen if one attacks these three algorithmic frameworks by harnessing Type-II-cynetbits-based QSEs. To answer that question, it seems crucial to consider that concerning their respective *inputs*, these three algorithms have only focused on *quantities* of bits of information and *not* on the meaning of those. However, an end-to-end-Type-*I*-pipeline could never know a priori whether a specific string of bits encodes Type-II-cynetbits-based QSEs or not. Thus, for all cases where the inputs to these respective three algorithms would match Type-II-cynetbits-based QSEs, comparative experimental runs would fail since being *inconsistent* with the assumption of Type I quantum supremacy. To recapitulate, for any of those three algorithms, it would hold that it is *impossible* for a Type *I* quantum computer to reliably process a respective *cynet-dependent* [1] version in a manner that is *qualitatively* superior to a Type I classical computer. For a simplified illustration, given each of the three algorithms, I provide *one* of the many possible examples for a *cynet-dependent* formulation:

1) *Cynet-Shor-algorithm:* To put it very simply, Shor's algorithm could be described as a function $f_{Shor}(N) = L_p$ taking as input an integer N and identifying the list L_p of its prime factors[3]. What I call the Cynet-Shor-algorithm is a function $f_{CynetShor}(f_{Type_{II}Cynetbits}(x)) = L_p$ that takes as input a function $f_{Type_{II}Cynetbits}(x) = N$ which given an explanatory IPS test x that hides a new EB (which is equivalent to the notion of Type-II-cynetbits-based QSEs), outputs the number-encoded combination N corresponding to the indexes of the blocks from that encrypted new EB.
2) *Cynet-Deutsch-Josza-algorithm:* To put it very simply, one could state that the Deutsch-Josza-algorithm $f_{DeutschJosza}(f_{Oracle})$ consists in making the binary choice of whether a black-box oracle $f_{Oracle} : \{0,1\}^n \rightarrow \{0,1\}$ that is promised to be either 1) constant or 2) balanced, is 1) or 2). In the first case, it would mean that all n inputs lead to a constant output (i.e. either all bits lead to 0 or all bits lead to 1). In the second case, half of the inputs would lead to the output 1 and the other half to the output 0. In short, with $n = 1$ it would be equivalent to identify whether $f_{Oracle}(0) = f_{Oracle}(1)$ (in which case it would be constant) or whether it rather holds that $f_{Oracle}(0) \neq f_{Oracle}(1)$ (in which case it would be balanced). What I term the Cynet-Deutsch-Josza algorithm is a function $f_{CynetDeutschJosza}(f_{Type_{II}Oracle}(x))$ whereby x is an explanatory IPS test that may or may not hide a new EB (which is equivalent to the notion of Type-II-cynetbits-based QSEs) and where it holds that the black-box oracle $f_{Type_{II}Oracle} : \{0,1\}^n \rightarrow \{0,1\}$ is constant if the n bits did *not* encrypt a new EB and balanced if the n bits *did* encrypt a new EB.
3) *Cynet-Grover-algorithm:* To put it very simply, Grover's algorithm can be seen as a function inversion algorithm. In brief, supposing a function $f(x) = y$ where *a unique input* produces a particular output, Grover's algorithm would be able to identify x given y. The underlying assumption here is that one has a database of N entries where $f(x) = 1$ only if the entry satistfies a certain criterium. The Cynet-Grover-algorithm could be equivalent to the task of identifying *the last block of a new EB* encrypted in an explanatory IPS test (which implies the notion of Type-II-cynetbits-based QSEs).

[3]Lists are ordered collections that allow duplicates. For instance, $f_{Shor}(40) = [2, 2, 2, 5]$.

IV. Synopsis

Due to cyborgnetic *emergence* phenomena, a shortcut for *understanding* is impossible. Simultaneously, it holds that in theory, the *creation* of any information form *except the creation of new EBs* can be forged. This asymmetry between the ability to create information of the form x and to understand x leads to various fundamental issues. It has already been elucidated that present-day Type I AI could be designed to output strings that are perceived by humans as representing EI and it is thus important to keep in mind that there is no underlying understanding by such classical Type *I* AIs in order to avoid honey mind traps [2]. In view of Section II and Section III, it becomes apparent why one must also avoid *quantum* honey mind traps. In this paper, I explained why *neither* a Type *I* classical *nor* a Type *I* quantum computer can be a universal computer. I introduced quantum sponge examples (QSEs) as adversarial attacks against Type I quantum computers. I harnessed Type-II-cynetbits-based QSEs (i.e. QSEs relying on encrypted new EBs) for thought experiments probing the vulnerability of Type I quantum AI.

I postulated that it is impossible for a Type I quantum computer to reliably solve: 1) the *Cynet-Shor-algorithm*, 2) the *Cynet-Deutsch-Josza-algorithm*, 3) the *Cynet-Grover-algorithm*. Also, for a Type I processor, the cynet-dependent versions are indistinguishable from the original ones. Thus, it even holds that it is impossible for a Type I quantum computer to *reliably* outcompete a Type I classical computer at: 1) Shor's algorithm, 2) Deutsch-Josza-algorithm, 3) Grover's algorithm. Only variety can destroy variety. Only entanglement can destroy entanglement. While Type II entanglement can decrypt Type I entanglement, it holds that Type I entanglement *cannot* decrypt Type II entanglement. Thus, only *cyborgnetic* entanglement can destroy arbitrary entanglement – at the cost of work via *understanding*. In brief, the best currently available EB states that instead of Type I quantum supremacy, what we have is *cynet-dependent universality* [1] via an entanglement of both Type I *and* Type *II* superinformation. A cyborgnet is able to self-program *if it decides so* [4]. A hereto willing cyborgnet could be a self-programmable universal computer [4]. *Only* a hereto willing cyborgnet could be a universal constructor [1].

References

[1] N.-M. Aliman, *Deepfake Quantum – A Serious Adversarial AI Game*. Kester, Nadisha-Marie, 2022.

[2] ——, *Cyborgnetics – The Type I vs. Type II Split*. Kester, Nadisha-Marie, 2021.

[3] V. Mavroeidis, K. Vishi, M. D. Zych, and A. Jøsang, "The impact of quantum computing on present cryptography," *arXiv preprint arXiv:1804.00200*, 2018.

[4] N.-M. Aliman, *Self-Folding Cynet Worlds – The Ladder and Tali's Paradox*. Kester, Nadisha-Marie, 2021.

[5] N. Liu and P. Wittek, "Vulnerability of quantum classification to adversarial perturbations," *Physical Review A*, vol. 101, no. 6, p. 062331, 2020.

[6] S. Lu, L.-M. Duan, and D.-L. Deng, "Quantum adversarial machine learning," *Physical Review Research*, vol. 2, no. 3, p. 033212, 2020.
[7] W. Ren, W. Li, S. Xu, K. Wang, W. Jiang, F. Jin, X. Zhu, J. Chen, Z. Song, P. Zhang *et al.*, "Experimental quantum adversarial learning with programmable superconducting qubits," *arXiv preprint arXiv:2204.01738*, 2022.
[8] X. Yuan, P. He, Q. Zhu, and X. Li, "Adversarial examples: Attacks and defenses for deep learning," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 9, pp. 2805–2824, 2019.
[9] A. Auffèves, "A short story of quantum and information thermodynamics," *SciPost Physics Lecture Notes*, p. 027, 2021.
[10] N.-M. Aliman, *Anagrammatic Singularities – The Cyborgnetic Clockmaker*. Kester, Nadisha-Marie, 2021.
[11] I. Shumailov, Y. Zhao, D. Bates, N. Papernot, R. Mullins, and R. Anderson, "Sponge examples: Energy-latency attacks on neural networks," in *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 2021, pp. 212–231.
[12] D. Deutsch and C. Marletto, "Constructor theory of information," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2174, p. 20140540, 2015.
[13] L. Bresque, P. A. Camati, S. Rogers, K. Murch, A. N. Jordan, and A. Auffèves, "Two-qubit engine fueled by entanglement and local measurements," *Physical Review Letters*, vol. 126, no. 12, p. 120605, 2021.
[14] L. Buffoni, A. Solfanelli, P. Verrucchi, A. Cuccoli, and M. Campisi, "Quantum measurement cooling," *Physical review letters*, vol. 122, no. 7, p. 070603, 2019.
[15] S. K. Manikandan, C. Elouard, K. W. Murch, A. Auffèves, and A. N. Jordan, "Efficiently fueling a quantum engine with incompatible measurements," *Physical Review E*, vol. 105, no. 4, p. 044137, 2022.
[16] G. Francica, F. C. Binder, G. Guarnieri, M. T. Mitchison, J. Goold, and F. Plastina, "Quantum coherence and ergotropy," *Physical Review Letters*, vol. 125, no. 18, p. 180603, 2020.
[17] L. Lami and B. Regula, "No second law of entanglement manipulation after all," *arXiv preprint arXiv:2111.02438*, 2021.
[18] D. Frederick, "Critique of Brian Earp's writing tips for philosophers," *Think*, vol. 20, no. 58, pp. 81–87, 2021.
[19] N.-M. Aliman, *Self-Climbing Cynet Tree – Hidden Entropy in The Biosphere*. Kester, Nadisha-Marie, 2022.
[20] M. Cortês, S. A. Kauffman, A. R. Liddle, and L. Smolin, "Biocosmology: Towards the birth of a new science," *arXiv preprint arXiv:2204.09378*, 2022.
[21] ——, "Biocosmology: Biology from a cosmological perspective," *arXiv preprint arXiv:2204.09379*, 2022.
[22] J. Degrave, F. Felici, J. Buchli, M. Neunert, B. Tracey, F. Carpanese, T. Ewalds, R. Hafner, A. Abdolmaleki, D. de Las Casas *et al.*, "Magnetic control of tokamak plasmas through deep reinforcement learning," *Nature*, vol. 602, no. 7897, pp. 414–419, 2022.
[23] L. Da Costa, P. Lanillos, N. Sajid, K. Friston, and S. Khan, "How active inference could help revolutionise robotics," *Entropy*, vol. 24, no. 3, p. 361, 2022.
[24] N.-M. Aliman, "Hybrid cognitive-affective Strategies for AI safety," Ph.D. dissertation, Utrecht University, 2020.
[25] ——, *Deepfake Science – Crafting Epistemic Defenses*. Kester, Nadisha-Marie, 2022.
[26] D. Blackiston, E. Lederer, S. Kriegman, S. Garnier, J. Bongard, and M. Levin, "A cellular platform for the development of synthetic living machines," *Science Robotics*, vol. 6, no. 52, p. eabf1571, 2021.
[27] M. R. Ebrahimkhani and M. Levin, "Synthetic living machines: A new window on life," *Iscience*, p. 102505, 2021.
[28] S. Kauffman, "Is There a 4th Law for Non-Ergodic Systems That Do Work To Construct Their Expanding Phase Space?" *arXiv preprint arXiv:2205.09762*, 2022.
[29] N.-M. Aliman, "Endnotes UEF," https://nadishamarie.jimdo.com/uef/, 2022, Online; accessed 01-June-2022.
[30] R. Noble and D. Noble, "Was the watchmaker blind? Or was she one-eyed?" *Biology*, vol. 6, no. 4, p. 47, 2017.
[31] D. Noble, "The illusions of the modern synthesis," *Biosemiotics*, pp. 1–20, 2021.
[32] A. Johanson, "Modern topology and Peirce's theory of the continuum," *Transactions of the Charles S. Peirce Society*, vol. 37, no. 1, pp. 1–12, 2001.
[33] S. Kauffman and A. Roli, "The world is not a theorem," *Entropy*, vol. 23, no. 11, p. 1467, 2021.
[34] S. A. Kauffman and A. Roli, "The Third Transition in Science: Beyond Newton and Quantum Mechanics–A Statistical Mechanics of Emergence," *arXiv preprint arXiv:2106.15271*, 2021.
[35] D. Aerts and L. Beltran, "Are Words the Quanta of Human Language? Extending the Domain of Quantum Cognition," *Entropy*, vol. 24, no. 1, p. 6, 2022.
[36] N.-M. Aliman, "Explanatory Blockchain Forgery?" https://www.nadishamarie.jimdo.com/cyborgnetics/, 2020, Online; accessed 01-September-2021.
[37] G. Aubrun, L. Lami, C. Palazuelos, and M. Plávala, "Entanglement and superposition are equivalent concepts in any physical theory," *Physical Review Letters*, vol. 128, no. 16, p. 160402, 2022.
[38] Y. Kim, F. Bertagna, E. M. D'Souza, D. J. Heyes, L. O. Johannissen, E. T. Nery, A. Pantelias, A. Sanchez-Pedreño Jimenez, L. Slocombe, M. G. Spencer *et al.*, "Quantum biology: An update and perspective," *Quantum Reports*, vol. 3, no. 1, pp. 80–126, 2021.
[39] C. Fields, K. Friston, J. F. Glazebrook, and M. Levin, "A free energy principle for generic quantum systems," *Progress in Biophysics and Molecular Biology*, 2022.
[40] L. Slocombe, M. Sacchi, and J. Al-Khalili, "An open quantum systems approach to proton tunnelling in DNA," *Communications Physics*, vol. 5, no. 1, pp. 1–9, 2022.
[41] J. Ramsay and D. R. Kattnig, "Radical triads, not pairs, may explain effects of hypomagnetic fields on neurogenesis," *arXiv preprint arXiv:2206.08192*, 2022.
[42] H. Zadeh-Haghighi and C. Simon, "Magnetic field effects in biology from the perspective of the radical pair mechanism," *arXiv preprint arXiv:2204.09147*, 2022.
[43] B. Leberecht, D. Kobylkov, T. Karwinkel, S. Döge, L. Burnus, S. Y. Wong, S. Apte, K. Haase, I. Musielak, R. Chetverikova *et al.*, "Broadband 75–85 MHz radiofrequency fields disrupt magnetic compass orientation in night-migratory songbirds consistent with a flavin-based radical pair magnetoreceptor," *Journal of Comparative Physiology A*, vol. 208, no. 1, pp. 97–106, 2022.
[44] L. D. Smith, F. T. Chowdhury, I. Peasgood, N. Dawkins, and D. R. Kattnig, "Driven spin dynamics enhances cryptochrome magnetoreception: Towards live quantum sensing," *arXiv preprint arXiv:2206.07355*, 2022.
[45] J. Xu, L. E. Jarocha, T. Zollitsch, M. Konowalczyk, K. B. Henbest, S. Richert, M. J. Golesworthy, J. Schmidt, V. Déjean, D. J. Sowood *et al.*, "Magnetic sensitivity of cryptochrome 4 from a migratory songbird," *Nature*, vol. 594, no. 7864, pp. 535–540, 2021.
[46] K.-S. Chae, S.-C. Kim, H.-J. Kwon, and Y. Kim, "Human magnetic sense is mediated by a light and magnetic field resonance-dependent mechanism," *Scientific Reports*, vol. 12, no. 1, pp. 1–11, 2022.
[47] L. E. Foley, R. J. Gegear, and S. M. Reppert, "Human cryptochrome exhibits light-dependent magnetosensitivity," *Nature communications*, vol. 2, no. 1, pp. 1–3, 2011.
[48] R. Sarimov, V. Binhi, and V. Milyaev, "The influence of geomagnetic field compensation on human cognitive processes," *Biophysics*, vol. 53, no. 5, pp. 433–441, 2008.

APPENDIX A
EBs EXPLAINING EMERGENCE PHENOMENA OF EBs

In the AI field, some assume that Type I AI could learn any thinkable task and that all tasks can be solved on the basis of bits. I refer to this prevailing stance as *the reductionist paradigm*. In diverse past cyborgnetic books [4], [10], [19], I elucidated multiple facets on *why* an EB is more than the sum of its parts. On the whole, by focusing on the task of creating *new EBs*, cyborgnetics [2], cynet information theory [4] and cynet dynamics [1] refute the reductionist paradigm. This refutation can be complemented by various novel developments in physics and beyond. Overall, some of the currently best available EBs on that subject can be e.g. classified in three categories: 1) cosmological lines of reasoning, 2) mathematically focused analyses and 3) superinformation-related hypotheses. I very briefly introduce some key take-aways from each category.

A. Cosmological Perspective

When considering biocosmology [20], [21]– a framework very recently introduced by multiple known physicists – it becomes clear that most present-day AI (which is non-living Type I AI) including so-called intelligent systems may not exhibit requisite variety when compared to Type I and Type II *life* due to the immense space of degrees of freedom that living entities add to the universe as a whole. Already the degrees of freedom exhibited by stars[4] may not yet be attained by even the most advanced present-day non-living Type I AI. Obviously, all non-living Type I AI exhibits less degrees of freedom than the living and self-replicating xenobots [26], [27]. In addition, from the perspective of cynet information theory and its independent cosmologically-focused branch, while the set of possible functions for Type *I* life may indeed tend to grow steadily [20], may reach infinities and is unpredictable ahead of time [28] as explained by biocosmology [20], it is important to *additionally* consider the special case of Type *II* life. For a compressed summary on how biocosmology can be extended by applying a cyborgnetic lens, see [29]. Overall, following cynet dynamics [1], only a cyborgnet could be a universal constructor.

Indeed, Type II life, through the ability to consciously understand what a construct such as "possible functions" means, can consciously decide how to enact, enlarge but also to *reduce* those. Moroever, Type II life can also decide to behavioristically mimick selected trajectories and distributions. On the whole, Type II beings are *not* bound to biological imperatives or to the consideration of functions that are solely in the service of what is often described as "biological fitness". While biological entities can harness stochasticity at various levels [30], [31] leading to a partially sighted process including phenomena such as targeted mutations [30], Type II entities can craft EBs about what "stochasticity" signifies and can consciously employ it. Type II entities can literally even consciously manufacture selected mutations for their socio-culturally constructed goals. Moreover, voluntary suicide or the conscious destruction of the biosphere become possibilities. The result is that when trying to grasp something like the "number of possible functions" for Type II entities, one suddenly encounters an aggregate of abnumeral infinities – what Peirce called *supermultitudinous collection* [32].

B. Mathematical Perspective

Following Kauffman and Roli, the affordances that *living* entities enact in their biological milieu *cannot* be expressed via set theory [33]. They state that *"we can create no mathematical model of the diachronic evolution of the biosphere based on set theory"* [33]. One cannot mathematically predict those *ahead of time* [34]. Concerning Type II life, I postulate that one must strictly speaking consider the mentioned concept of a supermultitudinous collection – which is neither a set nor even a category from category theory. It is an ultra-dense condensate of genuine infinity of which there exists no higher order. As described by Peirce, the elements of such a collection are *not* points, but triadically *interdependent potentials* (see [32] for an in-depth explanation). In brief *"a supermultitudinous collection sticks together by logical necessity. Its constituent individuals are no longer distinct and independent subjects. They have no existence [...] except in their relation to one another"* [32]. I assume that it is for this reason that no mathematical formula can predict or postdict the creation of an unknown new EB and no formula can cover the entire potential of cyborgneticity.

C. Superinformation-Related Perspective

Interestingly, Aerts and Beltran [35] recently corroborated that new EI in the form of stories (such as Winnie The Pooh) can be interpreted as a special form of superinformation[5] since they were able to experimentally corroborate that – when directly compared to the classical Maxwell-Boltzmann statistics – *Bose-Einstein-statistics* represented a *superior* model for those texts. The authors elucidate that the latter may represent an explanation for the appearance of Zipf's law [35] known in computational linguistics. Prior to that, in the second cyborgnetic book [4], I postulated that specific new EBs (for instance anagrammatically encrypted ones and generally those intermingled in non-EB-like EI) are expressible as a new form of superinformation [2]. (i.e. more than assembled I pieces). In the third one [10], I generalized it to the statement that any *new* EB is a form of Type *II socio-psycho-techno-physical* superinformation while *new* non-EI-like LI or *new* non-EB-like EI can act as Type *I* socio-psycho-techno-physical superinformation – since it can be forged by Type *I* entities even though those do *not* understand it. (Note that since new EB forgery is impossible [36], new EBs represent a stronger form of superinformation that is only accessible to Type *II* entities [2].) Finally, what is conventionally described as quantum information can be described as a special case of *physical* Type *I* superinformation[6].

[4]Perhaps a hypothetical fictive future nuclear fusion reactor based on non-living Type I AI could reach this level. However, the application of non-living Type I AI to significantly improve nuclear fusion is currently only at the beginning with deep reinforcement learning [22]. Further, it is thinkable that active inference [23] could enhance the required non-living intelligent system – which may however come of the cost of predictability. What is more, also in this relevant context, one must consider context-dependent harm models such as augmented utilitarianism [24] and one must inject Type-II-ness for a cyborgnetic risk management applied to a COOCA-loop [25].

[5]Superinformation is a *scale-independent* term introduced in constructor theory of information [12]. Quantum information is only one special possible form of superinformation.

[6]When considering quantum information, it is important to keep in mind that it involves *mathematical* and thus substrate-independent formalisms and one must thus avoid the substrate-dependency-fallacy of mentally *a priori* reducing it to miniscule e.g. subatomic or atomic scales. What seems relevant to superinformation instantiating a non-classical paradigm are the notions of *entanglement*, *superposition* and *encryption* [37] – all of which are *not* a priori tied to a specific scale. In modern days, the discipline of quantum biology [38] gained momentum [39]. While only in its infancy, it already provided some experimental corroborations of quantum effects at many corresponding steps of the ladder: for instance at the level of DNA mutations [40], in cell-related oxidative stress mechanisms [41], [42], in living but non-conscious Type I entities such as plants or in conscious Type I life such as birds [43]–[45]. Human magnetosensitivity [46]–[48] (but so far *without* consciously accessible sense of it) and its conjectured link to spin dynamics [42] may thereby perhaps offer novel avenues for future yet unknown new EBs.