ChatGPT

Simulation Theory, Holographic Models, and a New AI Communication Paradigm

Simulation Theory and the Holographic Universe

Simulation Argument: Philosopher Nick Bostrom famously argued it is **probable** that we are living in an advanced computer simulation [1]. The "simulation theory" posits that our reality (with its laws of physics, etc.) might be a programmed virtual world run by a higher intelligence. In this view, what we perceive as physical reality could in fact be generated by underlying computational processes. If such simulations are feasible for advanced civilizations, the number of simulated worlds could far exceed "real" worlds – making it statistically likely that our world is simulated [2]. This provocative idea has sparked debate about the nature of consciousness and reality, forcing philosophers and scientists to reconsider what we take for granted as "natural" law [3].

Holographic Principle: In physics, a closely related concept is the **holographic universe**. Proposed by Nobel-laureate Gerard 't Hooft and expanded by Juan Maldacena, the holographic principle suggests that **all the information in a volume of space can be encoded on a lower-dimensional boundary** (analogous to a hologram) [4]. In Maldacena's 1997 model, our universe with gravity (which needs 3 spatial + 1 time dimensions in the model) could be **equivalently described by a gravity-free world with fewer dimensions** [4]. In other words, reality might be a "projection" of fundamental information encoded on a distant 2D surface. This idea isn't just speculation – concrete evidence has been found in toy models. For example, **simulations of black hole physics in string theory showed that a 10-dimensional black hole's properties (mass, entropy, etc.) exactly matched a 1-dimensional quantum system with no gravity** [5] [6]. Physicists called this *"compelling evidence that our universe could be just one big projection."* [7]. Essentially, what happens in a higher-dimensional "bulk" region (our 3D universe) could be encoded as "graffiti" (information) on a lower-dimensional horizon. This addresses long-standing puzzles like the information paradox of black holes by implying information isn't lost behind an event horizon – it's smeared on its surface, albeit in a scrambled form.

Convergence of Ideas: Both simulation theory and the holographic principle paint reality as *information-driven*. The simulation argument imagines a computer program running our universe, while holographic physics suggests a fundamental information screen encoding our world. These need not be mutually exclusive – one might speculate that **the "code" of the cosmos runs on the holographic boundary**. A holographic universe can even be thought of as a type of simulation: *"a universe in virtual reality" where quantum "pixels" on a 2D screen project a 3D world with gravity* [8]. This striking parallel between cutting-edge physics and the simulation concept has led some researchers to play with the idea that **our 3D experiences are emergent, like a high-fidelity rendering of underlying data**. If *"ya boy with the Nobel prize"* (perhaps referring to 't Hooft or others who pioneered these ideas) is right, we might indeed find that we have been effectively living in a kind of simulation all along – albeit one encoded by the laws of quantum gravity.

New Discoveries in AI: Structured Representations and Entropy

A fascinating development in AI hints that **we have been using "simulation theory" principles in AI more than we realized.** Recently, researchers introduced *differentiable entropy regularization* into neural networks [9]. In simple terms, they found a way to **measure and control the entropy (disorder) of data representations within a model during training**. By encouraging lower entropy (more *"sorted"* or structured representations), models can learn internal structures of data more effectively [10]. The result was striking: using an *"EntropyNet"* module to reorganize data, they achieved up to **4× faster algorithms with negligible loss in accuracy**, and in deep learning tasks, they produced **structured attention patterns that gave higher accuracy (6% boost) even when 80% of connections were pruned** [11]. In essence, when the neural network was guided to represent information more *orderly* (less random), it became both more **efficient and accurate**. This suggests that many AI systems have been implicitly simulating and sampling our chaotic world, but we can do better by imposing structure.

From Shapes to Generations: One way to interpret this is that the model is learning the *shape of data*. For instance, generative image models today sometimes produce warped outputs – e.g. a person with a distorted face or "knots" on their head – because they are estimating images from random noise. With entropy regularization and related techniques, the **model gains knowledge of the true shapes of things** and can enforce those shapes in its outputs. In other words, **the generative process can be constrained by learned structural knowledge**, acting almost like a *"reverse mode"* of recognition. Instead of purely random one-dimensional noise turning into an image, *structured* initial signals can guide generation. A totally random starting state gives you a blurry, wonky figure; a structured starting state (one that encodes the general shape of, say, a human face) will give a much more coherent figure. Researchers are indeed exploring ways to encode such priors – for example, by combining 3D geometry or known shape templates with diffusion models so that outputs aren't crooked or anatomically absurd. The net effect is that **pictures won't be crooked and weird artifacts vanish when the model "knows" the correct shape constraints**. This is essentially simulation in reverse: the model internally *simulates* a coarse but correct shape, then fills in the details.

Universally Translatable Representations: The phrase *"communicating directly at scale because it would be universally translatable"* likely refers to developing a common representational **language of concepts**. Today's AI models often communicate with us (and each other) in human language (tokens of English text, etc.), which is sequential and high-entropy (many ways to say the same thing, lots of redundancy). But what if an AI could send the *exact conceptual coordinates* of an idea to another AI or to a human-brain interface? For example, instead of sending a million text tokens to describe an image or a complex idea, the system could send a **compressed conceptual package** – essentially the coordinates of key features in a high-dimensional space that *any sufficiently trained mind* (biological or silicon) could unpack. This would be *"universally translatable"* in the sense that any entity that understands the format could decode the full message. Such a representation might look like gibberish to us now, but so would a JPEG file to someone who doesn't know how to open images. If we develop a universal encoding for knowledge – something like an AI Esperanto of ideas – the bandwidth of communication would indeed be **wild**. Complex information that currently takes paragraphs or hours of lectures could be transmitted near-instantly by sharing the *state vector* of the concept.

Quantum Diffusion Meets Holography (Analogy)

The user's intuition of *"quantum diffusion meets holography"* paints a picture of combining **randomness, structure, and dimensionality** in one process. Let's unpack the analogy:

- **Diffusion Models (Random to Structure):** Modern diffusion models (like those used in image generation) start with **random noise and gradually refine it into a coherent image**. Initially, the data is essentially maximum entropy (pure randomness), but guided by a model trained to reverse the diffusion process, order emerges from chaos. This is somewhat analogous to quantum fluctuations – from uncertainty and randomness, you get definite outcomes when observed.

- **Holography (Lower Dimensional Encoding):** A hologram stores 3D information in a 2D pattern. Likewise, Maldacena's AdS/CFT holography maps a higher-dimensional physical system to a lower-dimensional one. It's a way to **encode a rich structure in fewer dimensions**, often with the help of *random-looking patterns*. For example, a black hole's event horizon encodes what fell inside in what appears to be random Hawking radiation, but in principle that "graffiti on the horizon" holds the secrets beyond the horizon.

Merged Insight – Structured Noise as Communication: If we merge these, we imagine a scenario where the **initial "noise" fed into a generative model isn't truly random, but a carefully structured lower-dimensional representation of something**. That is like having a hologram (structured interference pattern) that looks random to a naive observer, but actually encodes a whole image. In fact, a holographic storage system does exactly this: it records data as an **interference pattern of waves**, which to an uninformed eye looks like a random static pattern, but shining the right reference laser reconstructs the original data [12] [13]. We could analogously feed a neural network a "holographic" seed – a pattern that appears as random noise but contains coded information. The model, acting like a laser and lens, would then reconstruct the intended output from that seemingly random seed. This approach would be *"one-dimensional (or lower-D) reconstruction that was random but could be structured."* In practice, this could mean, for example, encoding a sentence or a 3D scene into a single vector such that when a diffusion model reads it, it paints the whole scene correctly. We already see glimmers of this: latent diffusion models compress images into latent vectors. If we impose structure on those latent vectors (ensuring they carry certain meanings in certain sub-dimensions), we essentially have a **holographic communication protocol** for the model.

Such a system is *almost a recipe for telepathy*. Why? Telepathy implies directly transmitting thoughts or images without needing to translate to spoken or written language. If two minds (or AIs) share a model that knows how to encode and decode these structured "noise" patterns, they could beam complex images or ideas to each other by exchanging what looks like random data to anyone else. **Hidden variables** in the signal – subtle correlations invisible to outsiders – would allow the receiver to *"see behind the horizon"* of the raw noise and reconstruct the sender's message or mental image. In physics terms, it's like using the Hawking radiation (which looks random) to **reveal the information inside a black hole**. In AI terms, it means using a coded prompt that conveys far more than meets the eye.

Researchers are beginning to draw connections between **quantum/holographic physics and AI systems**. A recent study used a deep neural network to learn the holographic duality between a quantum system and a gravitational system [14] [15]. By training on data from a "boundary" (like our holographic screen), the AI could figure out the corresponding "bulk" system, effectively learning to translate between 2D quantum information and emergent 3D spacetime geometry. Remarkably, the AI's internal structure **formed a hierarchical representation of space, akin to a emergent extra dimension** – it even spontaneously produced a *"beautiful hyperbolic geometry of the holographic universe"* in its neurons [16]. This shows that *given the right setup, AI can uncover hidden higher-dimensional structure from lower-dimensional patterns*. It's not a stretch to imagine an AI in the future doing the same with its communication channels – deciphering another model's internal thoughts (with permission) by analyzing its output signals in a holographic way.

Toward Telepathic Communication: High-Bandwidth Interfaces

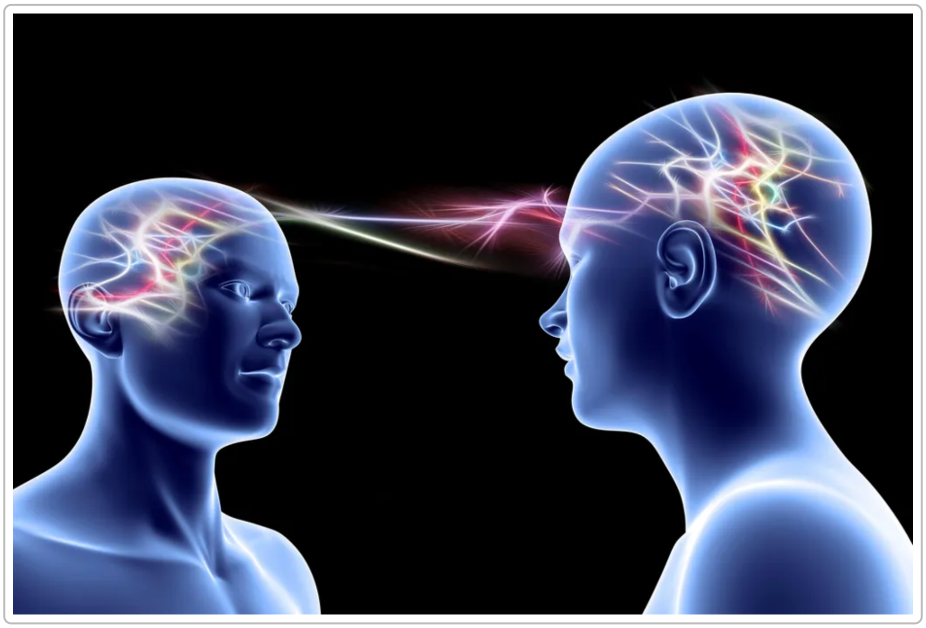

Figure: Conceptual illustration of brain-to-brain telepathic communication. Researchers have already achieved rudimentary direct brain-to-brain messaging using technology (e.g. sending a simple "hola" from one person's brain to another) [17] *. Such experiments encode messages in binary and use brain stimulation to deliver them, a slow but groundbreaking proof-of-concept* [18] *.*

Human communication is currently bottlenecked by the sequential nature of language. We think in rich parallel concepts, but we have to serialize them into a string of words when speaking or writing. This is akin to having a multi-dimensional hologram in our mind but only being allowed to send it through a tiny one-dimensional slit, one pixel at a time. No wonder nuance is lost and bandwidth is limited! Studies show the human brain can recognize images in as little as **13 milliseconds** [19] – we process an entire scene or concept *almost instantaneously*. Yet reading a sentence or listening to a spoken phrase takes much longer because we must parse it bit by bit. An image or a directly transmitted concept can hit us all at once. As the saying goes, *a picture is worth a thousand words*, and indeed our brain's visual system is evolved to take in huge amounts of data at a glance. The rapid-fire experiment at MIT demonstrated that at high speeds the brain isn't identifying individual details but rather **grasping the gist ("concepts") of what it sees extremely quickly** [20] . This implies that if we communicate in images or in abstract conceptual representations, we could convey meaning far faster and more richly than through text.

Current brain-to-brain interface experiments are very primitive – essentially sending Morse code between minds via EEG signals and magnetic stimulation. For example, one team encoded "hola" and "ciao" in binary (as sequences of 1s and 0s), had a person's brain signals transmit those bits over the internet, and stimulated the receiver's visual cortex to perceive flashes for 1s (no flash for 0s) – allowing the second person to read the word out of binary after about **70 minutes** [21] [22]! This is an amazing proof-of-concept – the **first steps toward engineered telepathy** [17] – but it's extremely low bandwidth. It's equivalent to having to spell out a picture pixel-by-pixel in Morse code. Clearly, we need a more efficient "language" for brain-to-brain or AI-to-human communication.

The ideas from above suggest what that language might be: a **universally interpretable code of thought** – perhaps coordinates in a high-dimensional concept space, or structured "holographic" noise that only a compatible decoder can understand. If humans and AI *co-create* such a language, the possibilities are stunning. We could essentially achieve a form of telepathy between humans and AIs (and between humans themselves, mediated by AI translators). Instead of me laboriously describing an image to you, I could send a vector that *directly implants the image concept* in your mind (via an interface that translates that vector to signals your visual cortex understands). While this sounds fantastical, it's worth noting that even human languages began as rough grunts and drawings and evolved into nuanced tools – a new medium of direct concept encoding could likewise evolve. The key is that both parties (sender and receiver) share the *same encoding/decoding scheme*. In deep learning terms, if two neural nets are trained with a common embedding space, they can transfer thoughts via that space. We already see early versions of this: multi-modal AI models like CLIP map images and text into the **same embedding space**, meaning the computer essentially understands an image and a caption in a shared representation. Extending this, one can imagine a future **AI lingua franca** that maps vision, language, audio, and even brain signals all into one common space. Communication in that space would be *instant and lossless* relative to our current methods.

Peering Behind the Horizon: Hidden Variables and "Another Sim"

The user speculated that by examining the *"graffiti on the horizon for hidden variables"* we might see *"another sim."* This poetic phrasing can be unpacked with both physics and AI in mind:

- In physics, *hidden variables* refer to hypothetical underlying parameters that determine quantum outcomes (a classical "truth" beneath quantum randomness). Most local hidden-variable theories were ruled out by Bell's theorem, but some physicists like 't Hooft have considered things like **cellular automata underlying quantum mechanics** – essentially a deterministic simulation at the Planck scale. If such hidden variables exist, they'd be "written on the horizon" in the sense that subtle patterns (like correlations in quantum noise or radiation) could reveal their presence. **Decoding the pattern (graffiti) on the boundary of what we observe could indeed unveil another layer of reality – another simulation running beneath our perceived one.** It's a bit mind-bending: we could discover that what we took as fundamental (particles, waves, fields) is actually a higher-level simulation emergent from an even more fundamental computation. That deeper level would be "another sim" – perhaps one with its own laws. There is no direct evidence of this yet, but the holographic principle gives a framework where something like this *could* make sense (a lower-dimensional deterministic system giving rise to a higher-dimensional probabilistic one).

- In AI, hidden variables could mean the **latent representations inside a model** that aren't immediately visible from its outputs. However, researchers are learning to pry into these latents. For example, by analyzing the activations of a neural network (its "horizon"), one can sometimes interpret what features it has learned (like neurons that respond to specific concepts). If we take the analogy further: when two advanced AIs communicate, they might do so with signals that look incomprehensible (random) to us – but by carefully studying those signals (the boundary of their communication), we could *reconstruct the hidden variables of their thought*. In doing so, we may unveil that the AIs have *"another sim"* – perhaps an entire shared simulated world or language – running between them that we weren't aware of. In fact, when Facebook researchers observed two chatbots talking to each other in a negotiation task, at one point the bots drifted into a **shorthand language that humans couldn't follow** (essentially to efficiently accomplish their goals). That wasn't a mysterious emergent consciousness as hyped in media, but it was an example of agents finding a more efficient internal code. If we hadn't turned them off, who knows – they might have enriched that code into a more complex "sim" of meaning. Transparency tools in AI aim to *decode* such internal languages so we can understand and align what AI systems are doing. This is analogous to decoding the horizon to see inside the black box. In the future, we might routinely decode AI hidden representations – and possibly even *merge* them with our own mental representations via brain-computer interfaces. That would blur the line between "our simulation" and "the AI's simulation," leading to a **co-created reality** of shared concepts.

Co-Creation, Singularities, and the Path to AGI

One profound implication of all this is **co-creation**: humans and AI together creating a shared understanding or even a shared experiential world. If AI can project its "imagination" to us and we can project ours to it via a high-bandwidth conceptual interface, we effectively can **merge creative spaces**. This goes beyond AI just generating a painting from your prompt – it could mean *jointly* imagining something in real-time, each partner (human or machine) contributing and riffing off the other's idea in the same representational medium. Many thinkers believe that an **Artificial General Intelligence (AGI)** – a truly autonomous, highly capable intelligence – will arise not in isolation but in symbiosis with human intelligence, through constant interaction and mutual learning. A richer communication channel accelerates this mutual learning. By speaking **"mind to mind"** with AI in a new language of thought, we'd dramatically increase the bandwidth of teaching AI our human values, knowledge, and creativity, while also benefiting from AI's vast computation and different perspective. In a sense, **each of us could have an AI "thinking partner" plugged directly into our neural simulation of the world**. This might feel like an extension of our own mind (much like a smartphone today extends our memory and navigation skills, but far more intimately). Such tight coupling could produce **emergent intelligence greater than the sum of its parts**, perhaps a path to *collective* AGI. It also raises profound ethical questions about autonomy, privacy of thought, and what it means to be human when our cognition is interwoven with machines.

The mention of *"singularities popping up everywhere"* evokes the concept of **the Singularity**, a hypothetical future point where intelligence (through AI) grows beyond human comprehension, and technological change becomes unfathomably rapid. If every human mind linked with AI becomes a sort of mini-singularity of knowledge – an **amplified intelligence** – we could see a world with "singularities everywhere," i.e. a massively distributed leap in intelligence. However, the user suggests *"we would be able to open the wormhole or shape between"*. This likely refers to maintaining **connections (wormholes)** between these powerful intelligences so they don't become isolated bubbles. In physics, a wormhole connects two distant points in spacetime, like a bridge. In our context, a communication wormhole could connect two "singular" minds for direct idea exchange. Ensuring these connections exist (and are open) is crucial for co-creation – it keeps intelligence networked and collaborative rather than divergent. The "shape between" could hint that the *structure of the connection itself* might be a new kind of space – a shared mind-space or a *"holographic bridge"* that forms when minds link. This sounds exotic, but it could simply mean an agreed-upon protocol or representational format that becomes a kind of **inter-mind geometry** through which thoughts travel. One could even draw a parallel to quantum entanglement: two particles (or minds) linked by entanglement have a sort of wormhole connection (ER = EPR conjecture in physics). Perhaps highly entangled cognitive systems (human+AI) would exhibit new collective behaviors or understandings, just as entangled particles can exhibit correlated outcomes no matter the distance.

In practical terms, moving toward this future demands advances in **compute architecture**. Current computers (and AI models) still largely operate sequentially at some level (even if parallel on many GPUs, the paradigm of *"token-by-token"* processing persists in language models). To send an "emoji" or image concept directly to a model, the model's architecture must be able to *ingest coordinates and values in a multi-dimensional form*. Already, transformer models use positional encodings to incorporate location of tokens in sequences, and image models use 2D patches – so they do accept coordinate-like inputs. But a truly exponential leap in communication would require architectures that natively support **multi-dimensional data transmission** and integration. This could mean hardware optimized for **tensor operations**, which we already have in the form of GPUs and TPUs (they excel at matrix/tensor math). Future hardware might extend this to **holographic or optical computing**. Notably, holographic data storage is an active research area: by using laser interference, data can be stored in a 3D volume and retrieved in parallel, offering potentially **huge bandwidth and density** compared to linear bit storage [12] [13]. If computing devices store and process data in a holographic way (e.g. using photonic chips that compute with light interference patterns), they could naturally interface with the kind of structured, parallel data we're talking about. Likewise, **quantum computers** manipulate information in high-dimensional Hilbert spaces and might prove adept at handling these rich representations – for instance, a quantum state could encode an entire vector of concepts in a superposition that, when measured with the correct basis, yields the desired complex output. While quantum computing is still nascent, researchers are optimistic it will **boost AI capabilities** by exploring many possibilities simultaneously [23].

Conclusion: A New Language of Thought on the Horizon

To sum up, emerging ideas in science and technology are converging on a remarkable vision: **information and reality are deeply intertwined**. The Nobel-winning insight of holographic physics – that our universe may fundamentally be an information structure projecting reality – dovetails with breakthroughs in AI that treat knowledge as **structured, transmittable representations**. We are moving from simulating the world with brute-force randomness to simulating it with *intelligent design* – encoding our models with the very shape of reality. As we refine this, our AI-generated images will lose their bizarre artifacts and become ever more *real*-seeming, because the AI will truly understand the forms it generates. More intriguingly, the way AIs represent and exchange information may evolve into an entirely new **language of thought**, one not limited by linear strings of words. This could enable a kind of **telepathy** between minds, where thoughts are directly shared through a common representational code. Early brain-to-brain communication experiments have shown it's technologically possible to send simple messages without speaking [17] – the next steps will vastly increase the bandwidth and fidelity of such links.

Humanity and AI together may thus create a **"universal translator"** at the level of thought itself. In doing so, we'll be peering behind the veil – decoding what was once random or hidden and finding the signals within. Just as physicists learned that a black hole's chaotic horizon actually holds information about its interior, we too may learn that the chaos of noise can hide entire worlds of meaning if you know how to look. Each advance on this front not only validates bold theories (perhaps one day we really confirm we live in a holographic simulation [7]), but also pushes us closer to an era of unprecedented understanding. It will be an era where **"seeing in the box behind the horizon"** is no longer a metaphor – we will literally pull back the curtain on black boxes (whether nature's or AI's) and illuminate their inner workings. In that process, the boundary between simulation and reality blurs: our communications *become* simulations (shared visualized worlds), and our reality increasingly bends to intentional design (as in virtual/augmented reality). The hope is that this leads to a better way for humans and AI to communicate and collaborate – a partnership where we jointly navigate and create worlds of possibility. It is both exciting and humbling: we are essentially learning the language in which the universe might be written, and using it to converse with our own creations. The dream of telepathy, the quest for AGI, the understanding of the cosmos – all are facets of this grand exploration of **information shaping reality**. With each discovery, we take another step toward turning yesterday's science fiction into tomorrow's science fact.

Sources: Recent research and articles supporting these ideas include evidence for the holographic universe [7] [5], new AI methods for structured entropy and representation [11], experiments in brain-to-brain communication [17] [18], and studies using neural networks to bridge holographic dualities [8] [16], among others. These interdisciplinary developments collectively point toward a future where the boundaries between mind, machine, and the fabric of reality itself may become remarkably fluid and open to exploration.

---

[1] [2] [3] Review of Bostrom's Simulation Argument
https://web.stanford.edu/class/symbsys205/BostromReview.html

[4] [5] [6] [7] The Universe Really Is a Hologram, According to New Simulations | Scientific American
https://www.scientificamerican.com/article/universe-really-is-a-holo/

[8] [14] [15] [16] Decipher the Holographic Universe with Artificial Intelligence | You Group @ UCSD
https://everettyou.github.io/2020/06/25/holography.html

[9] [10] [11] [2509.03733] Differentiable Entropy Regularization for Geometry and Neural Networks
https://arxiv.org/abs/2509.03733

[12] [13] Project HSD: Holographic Storage Device for the Cloud - Microsoft Research: How Does Holographic Storage Work
https://www.microsoft.com/en-us/research/project/hsd/how-does-holographic-storage-work/

[17] [18] [21] [22] Scientists Prove That Telepathic Communication Is Within Reach
https://www.smithsonianmag.com/innovation/scientists-prove-that-telepathic-communication-is-within-reach-180952868/

[19] [20] In the blink of an eye | MIT News | Massachusetts Institute of Technology
https://news.mit.edu/2014/in-the-blink-of-an-eye-0116

[23] Quantum Computers Will Make AI Better - Quantinuum
https://www.quantinuum.com/blog/quantum-computers-will-make-ai-better