The Cometary Biosphere and the Origin of Life

Robert B. Sheldon

NASA/MSFC/NSSTC/VP62, 320 Sparkman Dr, Huntsville, AL, USA

ABSTRACT

The Origin-of-Life (OOL) is defined as an information threshold and compared to the Shannon information of the universe. It is shown that the information content of a minimally viable cell must be greater than the capabilities of the universe to calculate with a random search, and must therefore include coherence. Since No-Free-Lunch theorems argue that there are no better algorithms than random searches, we eliminate several alternate theories of OOL that rely on "smart" algorithms, including the anti-entropic "luck" solution. Then high negentropy states can only be achieved by coherent addition of pre-existing negentropy via some low-entropy mechanism. Since most cosmologists believe information is conserved, it is shown that the addition of information corresponds to a flow of information through Fourier space from large to small scales. The requirements on the information "adder" for low temporal entropy, high spatial coherence, rapid coherent addition, and dense Fourier space flow, are shown to be met by comets. We close with a speculation that the fractal dimension of the galactic matter distributed through the cosmos may reveal the details of a dark matter origin in comets.

1. INTRODUCTION

The chance origin-of-life is a necessary precursor to Darwinian evolution but one whose plausibility has diminished in the 150 years since Darwin.[1] Some would argue that it never was part of Darwin's theory, since its principle mechanism for change, natural selection, cannot operate abiotically as OOL must. Nevertheless, Darwin included warm ponds in his monumental work, and insisted that chance OOL must exist given enough time.[2] As proof that chance OOL is part of evolutionary theory today, we note that when panspermia theorists insist that there is no need for evolution–rather life came from other locations and merely gave the appearance of evolution–the major criticism they receive is that panspermia theories lack an explanation for OOL and are therefore NOT a replacement for evolution. So this criticism reveals that OOL is very much a part of Darwinian evolutionary thought, and accordingly panspermia theories must also include it.

Since OOL is important for both Darwinian and panspermia theories, and because it is now possible to create "synthetic" life from machine-made DNA, OOL has become a topic of even greater interest. We address OOL by integrating it with transport, with panspermia. The result is a theory that is much broader than panspermia, which only argues for travel from planet to planet, while we argue for travel from star to star making life on planets a traffic accident in the vastly greater cometary biosphere.[3] Since we argue for an Earth biosphere embedded in a much greater cometary biosphere that is galaxy-wide if not cosmologically wide, we differentiate our theory from the more restricted panspermia, calling ours panzooia.[4] Panzooia is not simply a modified theory of OOL, it is an experimental hypothesis about comets that surprisingly had implications for OOL.

Paths to Panzooia There are several ways to present the relationship between comets and OOL, each beginning with a central observation and converging to comets. One starting point is to examine Darwin's concept of OOL in a "warm pond"[1] and show that the exergy,[5] or Gibb's free energy,[6] is not favorable for ponds or hot geothermal vents,[7] or gas-giant water planets,[8] but is favorable for comets. Another approach is to discuss the way in which coherence is required for sufficient information storage, which heat and entropy conspire to destroy, but which frozen comets preserve. Yet another is to look at addition of information involved in stepwise change (aka evolution), where the "adding machine" must itself be a coherent system that does not increase the entropy by "mixing", which is true of comet accretion. We combine all these threads to demonstrate that comets provide a solution to exergy, to coherence, to addition, and as a bonus, to the dark matter problem. But our first task is to define the problem in a quantitative way.

Further author information: E-mail: rbs@rbsp.info, Telephone: 1 256 653 8592

It bears repeating that science lives in a quantitative realm, whereas myth lives in a qualitative anarchy. By analogy to the cracking of the hieroglyphic code, there were many fanciful theories about the meaning of hieroglyphics until Thomas Young did a statistical analysis using the parallel Greek text found on the Rosetta stone.[9] In an analogous way, the OOL problem has been addressed with fanciful theories that are constructed top down, as complete explanations that qualitatively explain the phenomena, rather than systematically from the bottom up employing quantifiable measures of success or failure. That is, we think that the cracking of the OOL mystery will come through careful accounting of the necessary information involved and how it is retained or increased in the presence of thermodynamic decay. We address the nitty-gritty details of information bookkeeping lest we overlook some important step in our heuristic "ansatz" of how life began. While Young did not interpret the full text of the Rosetta stone, he did identify 200 words statistically, which was the method Champollion employed to finish the translation. Likewise, we do not propose to solve the OOL problem down to the distinct biochemical pathway and *in vitro* duplication of the event, but we do propose to find several information features of OOL that have been previously overlooked in the rush to guess an abiotic pathway to biological function.

In section 2, we define the OOL problem as an "Origin-of-Information" problem, and calculate a lower limit on the information present in a minimal genome. This approach bypasses having to speculate on chemical or biochemical processes that may or may not be involved, and simply calculates the informational content, which is indisputable. In section 3, we calculate the probability of OOL and discover that many computational methods are insufficient to achieve it. This includes not only local Shannon information, but evolutionary searches as well. In section 4 we show that the only method that computationally can achieve these magnitudes are those that are highly coherent, and merely combine or add pre-existing information. Since the addition of coherent information must itself be coherent, this puts stringent requirements on any putative process, which we list. Finally, in section 5, we show that the requirements of a coherent adder are satisfied by comets, which may even be primeval and at higher densities than presently modeled.

2. INFORMATION THRESHOLD OF THE ORIGIN-OF-LIFE

The OOL problem sounds trivial when one encounters it in the writings of Darwin, but in modern literature it has evolved to be a fearsome monster. While Darwin thought the cell was full of protoplasmic goop, we now know the cell to be composed of intricate nano-machines that are as thermodynamically far from equilibrium goo, as chicken soup is distinct from a squawking hen. One way to describe this difference, though by no means the final word in information theory, is to calculate the Shannon information contained by the two systems, and demonstrate the magnitude of the necessary dynamic information.

Then at the very least, OOL becomes the probability that we go from abiotic, low-information ingredients to biological, high-information machinery. This step involves the creation or injection of information, either through clever addition, spontaneous emergence, random search, or some combination of the three in the amount of the difference between the abiotic and biotic states. The amount of information needed to go from abiotic to biotic is then taken to be the probability of OOL for this situation. Multiplying this probability by all the places and times at which it might take place, then gives the global or cosmological probability for the origin of life. That is, multiplying the single event by the probabilistic resources of the universe gives the upper limit on the probability of the origin of life.

Misperceptions Before we calculate the probability of OOL, we need to address several statistical misconceptions about the calculation.

1) The first misconception is that because we see life around us, therefore the probability (integrated over sufficient space and time) must be one. However, Bayes Theorem the topic of conditional probability, demonstrating that the Probability of OOL given our present Data, $Prob(OOL|Data)$, is not the same thing as the probability of the Data, given the OOL, $Prob(Data|OOL)$.[10] Instead, they are related as follows using obvious abbreviations: $P(O|D) = P(D|O)P(D)/P(O)$; or, $P(O) = P(D|O)P(D)/P(O|D)$. The $P(D|O)$ is nearly 1.0, but given that panspermia seeks to explain the Data with some other mechanism than OOL, we might take that to be less than one. Likewise, $P(O|D)$ is very nearly zero, but may gain some probability from panspermia. But what exactly is $P(D)$? Multiverse theory would suggest that this number is close to zero, while others might

claim it to be nearly one. In conclusion, $P(O)$ is most definitely not 1.0, and may be much closer to 0.0, we just cannot use this method of argument to narrow the range without additional information and assumptions.

2) The second misconception is that a probability near zero is identically zero, so that even with infinite resources OOL has zero probability. This confuses an improbability with an impossibility, since infinity times an improbability is one, whereas infinity times an impossibility is still zero. The difference is profound, for an impossibility includes additional metaphysical assumptions. In this paper we take OOL to be an improbable, but not impossible event.

3) A third misconception is that the multiverse theory of infinite probabilistic resources renders any improbability to be a certainty (=1.0). We argue that not all infinities are alike, and without knowing the cardinality of the resources and the cardinality of the improbability, we cannot make this argument. (More precisely, an aleph-one infinitesimal times an aleph-null resource remains a zero probability.) Furthermore, many improbable monsters are called up by invoking infinite resources, each of which requires another assumption to tame, some of which are mutually contradictory, and all of which make infinite resources an unsatisfactory solution.

Having defended our calculation from these common dismissals, we continue with quest of putting a number on $P(O)$ or $P(O|D)$ by finding a reliable measure for the information in OOL.

3. CALCULATION OF COHERENT AND CONSERVED INFORMATION

The Distinction of life Since the advent of the 18th century, there have been debates about what distinguishes life from non-life. Early chemists separated organic from inorganic chemistry on the assumption that the chemistry of life was qualitatively different from the chemistry of non-life. The abiotic synthesis of urea shattered that paradigm, and today there is no qualitative barrier separating the two sub-disciplines, except for the extraordinary difficulty of abiotically making most of the proteins and nucleotides found in every cell. To give more precision to this concept of "difficulty", we argue that abiotic chemistry is driven by equilibrium or thermodynamic considerations, whereas biochemistry is far from equilibrium, which is why it is difficult to duplicate in a testtube. And the physical description of "non-equilibrium" thermodynamics is a low-entropy state, or what we will define later as Shannon information. Thus what separates life from non-life is neither the matter nor the energy, but the organization of the materials. As the late physicist, John Wheeler, said about his understanding of the world, "My career divides into thirds: everything is particles, everything is fields, and everything is information", which he applied to origins when he said "It from bit." Moving from physics to biology, biologist Manfred Eigen concluded that "the origin of life is the origin of information."

The significance of these statements is that there has been a late 20th century shift from a materialist to an immaterialist view of origins. Materialist John Paul Sartre in the early 20th century had claimed that "existence precedes essence", that complexity arose out of the pre-existing matter in some sort of emergent process. Wheeler and Eigen are concluding the opposite, that immaterial order preceded the matter that gave it expression. There has been a shift in viewing life as an intricate chemical reaction or a complicated nanomachine, to viewing life as the generation or propagation of information. Information reduces all these mechanical or chemical instances to an immaterial but quantifiable quantity. Thus no biochemist has the slightest hesitation to say that the information in the genome (+ epigenome, + proteome, etc) is the same as the information in a dynamic, living and metabolizing cell. Craig Ventner's experiment with "synthetic life" was an attempt to demonstrate that the DNA code was equivalent to life, and hence *in vitro* DNA is "synthetic life".[11]

A Measure for Life If the first qualitative step in distinguishing life from nonlife is in the quantity of information, then the second step is quantifying that information, quantifying the probability calculation by defining a measure or metric for it. One of the more popular measures is Shannon information, which Shannon defined as the negative of entropy, or "negentropy". Since entropy, S, is usually defined as the logarithm of the number of states of the system, $S = k\log(\Omega)$, then if every bit is independent of the other bits, if every state is bistable (on or off), we can count the number of states as 2^n whose logarithm is $n\log(2)$. This is the basis for Seth Lloyd's estimate of information contained in a computer, where a computer is an actualization of binary on/off states.[12] Alternatively, if the material actualization of states are "distinguishable", such as n ping-pong balls in the state lottery, then the number of states can be defined as all the permutations of n items, or $n!$, whose entropy is $log(n!) \sim n\log(n)$, which is sometimes called "information entropy". From quantum mechanics

Table 1. Showing five measures of complexity at three organizational scales.

Metric Scale	cellular	organismal	ecological
1. Size	Ostrich egg	Blue Whale	Earth
2. Diversity	organelles	cell types	species
3. Autonomy	modules	organs	ecologies
4. Machinery	expressed proteins	total proteins	bioengineered materials
5. Code length	euchromatin	genome	gene bank

considerations, there are severe limitations on "distinguishability" of both lightweight particles and very short intervals, which is why Lloyd uses $n\log(2)$ rather than $n\log(n)$ for computational information.

In order to apply Shannon information to OOL, we need a way to count states of the a/biotic system. Should we count the number of atoms, the mass, the number of nanomachines, the number of biological modules, the difference between kinds of modules, or perhaps the code defining those modules? Furthermore, should the fundamental unit of counting be cells, whole organisms, or even whole ecosystems? Since one can make an argument that multicellular organisms are more complex and more information rich than single celled organisms, or make the argument that every form of life is somewhat symbiotic and therefore the entire ecosystem complexity ought be calculated, we have provided a table of three levels of organization, and five different measures of information content. Each one has its advantages and disadvantages.

Using the method of examining the extreme cases to test validity, we argue that mass or number of atoms is a poor measure of complexity, or else ostrich eggs would be the most complex cell and blue whales the most complex animal, which seems intuitively incorrect. Likewise diversity is not the best judge of complexity, else the family of beetles is the most complex, with some 350,000 named species–70 times more species than all the mammal species on Earth. Likewise insects, the lineage to which beetles belong, include a million named species, the majority of all 1.8 million species scientists have ever described. (When JBS Haldane was asked what he could infer about the creator from his knowledge of biology, he quipped, "He has an inordinate fondness for beetles.") Therefore counting species is not the best way to identify complexity, in part because species can differ in trivial ways, and in part because species is not a well-defined quantity.[13] Likewise, counting the diversity of a cellular substructures (say a eukaryote versus a prokaryote) does not do justice to the complexity of the entire cell.

From Table 1, we see that rows 4 and 5 have both been used to classify the complexity of a cell. Row 4 looks at the "proteome", the number or interaction of the proteins needed for life, which Hoyle used to estimate the information content of a cell. Row 5 looks at the "genome" or the DNA instructions in a cell, which was the major justification for the human genome project that the genome was more basic than the proteome. (We could also have inserted a row 4.5 for the RNA-ome, which are intermediary molecules possessing some of the properties of the proteome and the genome, and which has the peculiar property of editing itself.) After hundreds of animals have had their genome transcribed, this claim to precedence is doubtful, but there is no controversy in saying that the information in the genome and epigenome and proteome is a subset of the information required for life.

The Calculation In calculating the minimum information content of life, almost all ignore the symbiotic relationships between cells, and the OOL probability is taken to be the information of the first cell, with multicellular organism and ecology expected to follow. In actuality, the first evidences of fossil life are stromatolites,[14] which are not only multicellular but also ecologically stratified communities. In partial resolution of this conundrum, we note that stromatolites are composed of prokaryotic cyanobacteria, which are viable as single cells but have the genes for loose multicellular organization as well as the unique ability to both photosynthesize and fix nitrogen. So while they can propagate as single cells, they are also capable of forming autonomous ecological communities–the mats that fossilize into stromatolites. It would be more consistent with these observations for OOL to calculate the information in a single cyanobacteria with minimal DNA length of about 1,300,000 basepairs,[15] but instead there is a concession to the "top-down" OOL theorists to use instead the minimal DNA length of any free-living cell.

Fred Hoyle attempted to derive the minimal information in a cell by considering the proteome as more fundamental than the DNA. This view is shared by "metabolism first" theorists, and argues that the DNA code could arise later after cellular replication gave a basis for natural selection to begin to work. A related theory is "RNA-world" in which RNA operates as both code and metabolizer. Both views have come under recent criticism as being too readily randomized by entropy,[16] but nevertheless remains a significant contender in the field of top-down theories. Since proteins or enzymes achieve their specific activity by their folding, the minimum length is on the order of 100-200 peptides, though Hoyle considered only the active region of 20 peptides, arguing that only one of 20! arrangements was usable.[17,18] Given the 20 kinds of peptides, and the 20 peptide length, Hoyle estimated Prob(one enzyme) = 10^{20} and therefore the 2000 enzymes thought to be needed for life gave a Prob(proteome) = $(10^{20})^{2000} = 10^{40,000}$ probability. Twenty years later Hoyle[19] updated their calculation by reducing the number of enzymes to 256, or Prob(proteome) = 10^{5120}. While this seems like a drastic reduction, they probably should have also updated their minimal enzyme length by increasing it to at least 62, the smallest enzyme known, which would have left their original calculation virtually the same.

Responding to the criticism that "metabolism-first" or even "RNA-world" is not capable of robust replication whereas DNA-world is, we can look for the minimal DNA genome length, and calculate the information in that. In 2010 when Ventor published his synthetic life, the shortest genome of any free-living (though parasitic) cell was that of *Mycoplasma genitalium* with a length of 582,970 basepairs. Since there are four base pairs to consider (ignoring the methylated versions of cytosine), we have Prob(genome)= $4^{582,970} = 10^{350,982}$. Corrections due to the 64 codons mapping to 20 amino acids, or the fact that there are some 500 proteins encoded that might be rearranged theoretically without loss of function, can reduce this to about $10^{250,000}$ though we stress these are all theoretical reductions, since in fact we know that duplicate codons and gene position do possess unique and necessary information.

Using modern genome manipulation techniques reported in Aug 2011,[20] a free-living bacteria, *Caulobacter crescentus*, had its genome shortened from 4 million basepairs to just 492,941 basepairs, making this the current record-holder for the shortest genome with $10^{296,780}$ bits of information. Whether this represents a continuing ability to reduce the genome, or an asymptotic size is unknown presently, but the paper does report many non-protein-coding, non-promoter sections of the *Caulobacter* genome that were still necessary for viability, demonstrating that many essential functions of DNA have yet to be understood. Despite our ignorance, the probabilities we are deriving are in the $10^{40,000}$ range for the proteome, to $10^{250,000}$ range for the genome with neither reducible to the 10^{50} range, which some have used as the threshold for something to happen in our universe as we demonstrate next.

The Information Deficit These numbers that we are calculating, $10^{40,000}$ or $10^{200,000}$, are so big it is hard to evaluate them, so one method is to compare them to calculation by the fastest computers physics allows. Let us assume that we have such an efficient computer that we can make a calculation with one electron, and that we can use every electron available in the visible universe, or about 10^{80}. Furthermore, if we run the computer at the fastest speed physics allows, where the shortest interval of time that has any meaning is the 10^{-43} Planck time, then there have been about 10^{60} intervals since the Big Bang. The product of those two is an upper limit on the number of calculations permitted by our universe, which Lloyd using more exact General Relativity considerations calculates to be 10^{120} operations on 10^{120} bits.[12] So we can subtract 120 from 200,000 giving only 199,880 orders of magnitude more computations than the universe allows.

One objection to this comparison, is that a random search by calculating all the permutations is not how evolution works, but that evolution is a "smart algorithm" that can zero in on the right sequence much faster. Is this possible? Even overlooking the observation that pre-biotic evolution, which is what OOL is, cannot make use of natural selection because there is no replication yet, there is yet a more fundamental problem. That is, for natural selection to select an optimal genome string, one has to find the genome string in the first place. And there are no "smarter algorithms" for finding genome strings, which do not also have encoded information about the desired string. If we have an ensemble of possible search algorithms without any information on the target string, none of them are on average any better than a random search, which is known in mathematics as the "No Free Lunch" theorem.[21]

Even worse, the space of search algorithms is even larger than the space of genome solutions, so it is not even possible to evolve a better search algorithm. Consider a 3x3 checkerboard with one checker somewhere on it.

We can classify efficient search algorithms as those that search every cell once. There are $9! = 362,000$ possible efficient search algorithms, whereas there are 9 possible configurations for the information. The search for a search is far worse than the search for just the information.[22] So this calculation of 10^{120} is a firm upper limit on the capabilities of uninformed searches in our universe. Random chance just does not have the probabilistic resources to find a living sequence.

Lady Luck If probabilities are against random chance, can we assume we had one truly lucky accident, or perhaps many smaller lucky accidents that add up to this number? If we take luck to be the discovery of Shannon information, then whether there are many steps or single steps, the amount of information involved remains the same. Since Shannon information is negative entropy, then this fixed amount of luck decreases the entropy by a fixed amount. For luck and entropy are opposites. Entropy is the iron rule of statistics, while luck is the violation of those statistics, of beating the odds, of winning the wager with the Devil. So in physicist terminology, the 2nd Law of Thermodynamics denies the validity or the ability to count on luck.[6]

In the words of cosmologist and physicist, Arthur Eddington, "If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations–then so much the worse for Maxwell's equations. If it is found to be contradicted by observation–well these experimentalists do bungle things sometimes. But if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation."[23]

For while one might get lucky with a few rolls of the die, on average there are more unlikely rolls, and by the time one has an Avogadro number of dice or dice rolls to average, the element of luck has vanished into insignificance. The more particles, the more rolls, the more solid the statistics. The 10^{80} particles and the 10^{60} rolls of dice in the universe already put stringent limits on the luck available in the universe, but this is not the only reason that Eddington is so negative about violating the 2nd Law of Thermodynamics. It is because microscopically we cannot tell if a movie of colliding gas molecules is running forward or backwards, but if we watch the entropy of the collection of molecules, if we watch the statistics, we can tell the direction, the arrow of time. A single molecule of water leaving the floor might be evaporating, but when a large number of water molecules on the floor gather into a heap and jump into the glass, we know the movie is running backwards, because such behavior is too lucky, it violates the statistical behavior of spilled liquids. Therefore, a theory that violates the 2nd law, is a theory running backwards in time, with causation running backwards too. This is how we define time, Eddington said, so it would be intensely embarrassing if we proposed a theory of reverse causation, of time flowing backwards.

Could it be a statistical fluctuation, a local increase of entropy while still being a global decrease, thereby avoiding the problem by not being a "closed" system? Not if it exceeds the computational resources of the entire universe, as we calculated above. There is not any known mechanism in the universe that can generate that sort of entropy imbalance. Even black holes are not capable of "destroying" entropy, according to leading cosmologists, who believe the universe is a closed system in which the amount of information and entropy is conserved (within the Lloyd limits listed earlier).[24]

For if OOL violates the Second Law of Thermodynamics, then it also violates causation, which was one of the assumptions in searching for a mechanism for OOL. With these sorts of numbers, there can be no appeal to luck. Following Eddington then, we must exclude all theories that have time flowing backward, that have an excessive amount of luck. Unfortunately, this means we must exclude many currently fashionable top-down theories. But in the name of science, we must not fudge the numbers, for the Hoyle put it, "numbers do not lie". We list four of the perennial favorites.

1. Emergence Can physical processes, like genetic algorithms or natural selection with mutations find the OOL genome string by searching randomly for it in the same way that computer simulations *Avida* or *Ev* evolve information? All of these programs have a "target" or fitness function, which is information encoded into the search. After accounting for that extra information, the search is no faster than random.[25] And as we discussed earlier, the necessary information is beyond reach of a random search. So we cannot start with a lucky, low-information state and evolve to a complex but less probable cell through any known random process, which can only add no more than 120 orders of magnitude when we need 200,000 more orders.

2. Multiverse If there are an infinity of universes popping out of the quantum vacuum, as Hawking and other theorists have proposed, could we avoid Eddington's restriction by preloading these universes with accidentally discovered information before their time's arrow began?[26] Despite the famous names associated with the multiverse hypothesis, it is not widely accepted by physicists because it makes so very many speculative assumptions. First, it assumes that our universe has a total energy of zero, which is controversial.[27] Second, it assumes that zero energy allows Big Bangs to spontaneously erupt in spacetime, which again has been experimentally challenged.[28, 29] Third, it assumes that each baby Big Bang started with slightly different initial conditions that can carry information, based on an even more speculative 11-dimensional "string landscape".[30] The net result of such rampant speculation is the final assumption that an infinity of such universes will provide adequate probabilistic resources to randomly find the living sequence of OOL.

It is this final assumption, however, that proves the undoing of the multiverse theory. For despite there being an infinite number of universes, there are many that cannot exist even in this infinite set. For example, time cannot flow backward in any of these universes, for in that case, the past big bang is a consequence of future universes, and "random chance" loses its meaning. So the requirement that time not flow backward means that the $10^{200,000}$ information must be encoded in the initial conditions–in the the position, velocity, and mass of each of the 10^{80} particles. But that means that on average each particle property must be specified to 199,900 decimal places, which is about 199,880 more decimal places than the Heisenberg Uncertainty principle allows, ($\Delta x \Delta mv \geq$ Planck's constant). Since there is no place to put this OOL information in the initial conditions and no way to generate the information later without reversing time, infinite universes do not help.

3. Infinite Space Can we be living in a universe that is much bigger than we can see, perhaps even an infinitely large universe that provides the probabilistic resources? This is again a controversial theory, since it invokes an asymmetry between space and time, with time to the Big Bang event horizon being 13.7 billion years but the radius of that Big Bang being essentially infinite. Even granting this assumption, however, the speed of light remains a constant, and therefore we can never access that infinity of space and its potentially infinite information. For the OOL problem is not only the huge amount of information, but that this information is condensed in a 2 micron cell. So the information resources available to that cell can only be collected from a sphere of 13.7 lightyears in radius, which is the visible universe. Therefore infinite space cannot produce infinite information density, or even the information in the first cell.

4. Infinite Time Since it seems the barrier to achieving the necessary information is the finite time of 13.7 billion years, can we somehow increase the amount of time available to infinity, can we eliminate the Big Bang or make it cyclic? Making the Big Bang cycle from Bang to Crunch and back to Bang endlessly cannot solve this problem, not only because the cycle should still have time's arrow and the 2nd Law, but also because it contains a finite amount of matter and therefore a finite amount of information encoded in that matter. Roger Penrose's cyclic cosmology is even more speculative, but he provides no indication how the "rescaling" of his cosmological cycle can increase the information content or even obey the 2nd Law.[31] On the other hand, removing the Big Bang to solve OOL introduces even more profound mysteries, which have been discussed ever since this "eternal" model was introduced by Democritus in the 5th century BC. Newton struggled with the problem of balancing an eternal universe against the force of gravity, and Olber's paradox asked why the night sky did not reveal either an infinite number of eternal stars, or as we understood nuclear fusion better, a black sky of infinitely many burnt out stars. But it was the discovery in 1961 of the predicted three-degree black body radiation that provided the final nail in the coffin of eternal or "steady state" models, such as those developed by Hoyle.[32] The astronomical paradigm shift was recorded by Robert Jastrow in his 1978 book, and despite the best efforts of Hoyle, the steady state model was never rehabilitated.[33]

In addition to cosmology, re-invoking eternal time produces problems for thermodynamics (which should be running down), mathematics (which abhors realized infinities), and especially theology (which has a vested interest in origins). Historically, centuries in which Epicurean eternities were implemented culturally have proven to be sociologically tumultuous, which suggests that the cost to society of denying a beginning is real. So from multiple arguments, we would suggest that this solution to the OOL information is at best problematical, and at worst inimical to science.

Conclusions Therefore it would appear that from considerations of information storage, information discovery, and information conservation, the OOL problem cannot be solved by invoking random chance.

4. COHERENCE AS THE SOLUTION TO THE DEFICIT

Coherence Defined But have we proved too much? It would appear that we have not only proved OOL impossible, but we have proven Hamlet could not have been written either! What is wrong with our calculation? In all our calculations we were assuming random causation, serial computation, locally isolated, Markovian processes. All this changes if we allow the parts to communicate with each other, if we allow parallel computations, long-range interactions in space (global) and in time (memory), which is the difference between combinations and permutations.

For example, in enumerative combinatorics there are 12 methods to bin N-particles in X-boxes: 3 conditions on X (any, surjective, injective) and 4 equivalence relations between X and N (equality, equality up to a permutation of N, equality up to a permutation of X, both N and x). In Lloyd's estimate of the universe-computer, he set N = number of atoms in universe, and X=N. Instead, we could quantize space for our boxes, which gives something X=10^{180} for boxes, and then N=10^{80}. This permits an even larger number than the 10^{120} but not enough for 10^{200000}. But an even greater increase is obtained if N and X are distinguishable, if the permutation order matters.

In layman's terms, if we have 5 boxes, and we can only record whether some box contains a ball or not, then we have 2 states for each box, or $2^5 = 32$ pieces of information in 5 bits. If, however, we label the boxes A, B, C, D, E, then we know even more about the system, and might rearrange the boxes in other ways: DECAB for example. The number of arrangements is $5! = 120$, or far more than simply 2^5. The difference arises because the boxes are distinguishable and hence rearrangements recognizable, unlike the previous case of 5 unlabeled boxes. The fact that each arrangement is distinguishable is related to coherence–a global snapshot that sees where everything is all at the same time. Note that in Einstein's spacetime we do not distinguish between spatial and temporal coherence–memory and patterns are equally informative. Mathematically, this coherence is counted by permutations, or $n!$, whereas the incoherent, local information, random situation is represented by 2^n. An approximation to $n! \sim n^n$, which for $n = 10^{80}$, is a lot bigger than 2^n. So the solution as to how our universe can contain massive information lies in its coherence. The information can be calculated in our universe if the calculation is coherent. But coherence requires a special sort of computer.

In order to illustrate the power of coherence to hold and process information, we need only look at the human brain. The density of neurons in our brain is about the same as all the primates, so the neuron count goes as the size. Since our brain is about twice the size of a gorilla, it would suggest we are only twice as smart, which is obviously not true. A closer look shows that each of our 30 billion neurons has 10,000 cross-connects, for a total of 10^{15} synapses. That's about 100 Terabytes, or 30 big hard disks. But experiments show that our brain stores more than that because it does not store memory in its synapses the way a hard disk stores photos, but each photo in our brain is distributed "holographically" over the whole region. That is, it is not the synapses, but the connections between the synapses that are important. If we have five synapses arranged in a loop, there are $5! = 120$ ways to arrange them versus the $2^5 = 32$ bits in a computer memory with 5 flipflops. Thus the brain is a coherent system, with not 10^{15} incoherent bits, but $(10^{15})^{10^{15}}$ coherent bits. This is the power of holographic or coherent memory storage. The key point is that information is not local, but global and distributed. Penrose argues that the coherence of our present universe provides something like $10^{10^{123}}$ bits of information that are unexplained by the random processes of the Big Bang.[34]

The Lloyd universe-computer was a serial computer, like all silicon chip processors, which is incapable of testing the 10^{200000} possible arrangements of the genome sequentially. A coherent quantum computer, however, can check them all simultaneously, if it had something like 1 million qubits for the calculation. The current limit on the maximum number of entangled qubits is about 20, a limit that doubles every 5 years or so, so we might be able to build such a computer by the year 2100, which of course, will reveal whether something more than DNA sequence is needed for viability. The point is that our universe is capable of bigger calculations than 10^{120}, but only if we employ coherent computations.

Coherence Observed How do we build a coherent computer, or how do we recognize a coherent calculation? Coherence is observed by calculating how much distant spatial or temporal locations behave or look like local ones. Mathematically, it is the autocorrelation function, or cross correlation function, which is implemented using the Fourier Transform. That is, if we want to see whether point A and point B are connected, we do a cross-correlation in space, or an auto-correlation in time. In both cases, we perform Fourier transforms on the data, and analyze the results in "transform space". This may seem mathematically abstract, but is precisely how our ears process sound for our brains to manipulate.

That is, when sound enters our ears, it is converted to pulses in neurons in the cochlea–a spiral organ that mechanically takes the Fourier transform of the waveform. These pulses then go to the brain, where phonemes, words, meaning, musical instruments are all extracted, before signals go to the larynx to convert digital frequencies back to analog waveforms. The power of hearing lies in the analysis of information in Fourier space. What is a long drawn out note on a flute in spacetime, becomes a very sharp frequency in Fourier space, and a sharp discrete pulse in spacetime is a blurry drawn-out blob in Fourier space. They are "complementary spaces". (A recent publication by Smolin and colleagues argues that Fourier or momentum space may be just as fundamental for the laws of physics as spacetime.[35]) Therefore all the information that is global, that is coherent, that is spread thinly all over real space, is discrete in Fourier space, and vice versa, so only by counting information in both spaces are we sure to have captured all the vital information.

We now develop some machinery for counting information, in the hopes of getting a more quantitative number for information than the 10^{120} used by Lloyd. This is a first attempt, and remains highly schematic, intended more to motivate than elucidate. Shannon information is defined as negentropy, which is the negative logarithm of the states of the system. $I = -S = -k\log(\Omega)$. Just as Smolin uses combined momentum and spacetime conservation laws, so we attempt to conserve information in both Fourier and spacetime. Working from Einstein's ansatz that they are aspects of the same 4-D reality, we generalize Shannon information to a 4-D space. Since information ought not to depend on reference frame, we write information as a scalar, $I = S^\alpha S_\alpha$. Since we also have information in Fourier space, and following Claude Shannon's example, we add them together $S^\alpha S_\alpha + FT^\alpha FT_\alpha$.[36] But the sum of logarithms is the logarithm of the product, so we attempt to fold these terms into a single invariant state vectors, $\prod_0^L (\Omega^\alpha\Omega_\alpha)_i$ where the i is an index indicating the segment of Fourier space being considered with $i = 0$ the ordinary spacetime, and $i > 0$ successively larger correlation distances up to the maximum extent of the universe $i = L$. The purpose of this is to show that most of the information in those 10^{200000} arrangements lies in the Fourier space terms, not in the local term. Then the paradox of OOL can be restated as observing Fourier information appear in real-space, without recognizing the Fourier space it came from.

Coherence Summed We have briefly alluded to the cosmologist's consensus that information is conserved in the universe, which means that as evolution progresses, it can only randomly generate the few 10^{120} bits/universe. Since this computed information is so insignificant compared to the pre-existing info, addition of pre-existing information is really the only way for information to accumulate for OOL. While there may be many adding steps that lead to OOL, each of those is at best the lossless adding of pre-existing information. This is supported by the oft-asserted claim that symbiosis is the driver of evolution, and illustrated in the example of eukaryotes being formed by the combination of an archaea and prokaryote.[37]

Then information before the addition must equal the information after, or, $I_p + I_a = I_e$, and when we analyze it, we see that coherence at large scales (the separate microbes) has now flowed into coherence at a small scale (the single eukaryote). Thus addition produces an increase in space-time information density by the method of flowing from larger to smaller Fourier scales. Evolution can then be seen as an information flow in Fourier space from big to small, concentrating the information in ever smaller bundles. Emergence, or the generation of complexity from simpler systems is denied, whereas as symbiosis or the addition of information is asserted.

Eukaryote evolution by symbiosis is still a controversial hypothesis, so it might be better to use a real-life laboratory experiment as a prototype "information addition." The experiment was conducted by Craig Venter, taking over 10 years and $40 million to accomplish. He asked if it was possible to create an organism with bits of added information that was completely "synthetic". Furthermore, he did not want to "splice in" the information like a virus, but he wanted to construct it from scratch, using a DNA gene-machine. Because if he could do it

Table 2. Inferred properties of coherent adders

Information Limitations	Coherence Spatial	Temporal	Adder Spatial	Adder Temporal
a) OOL cannot be random	√			
b) OOL cannot emerge		√		
c) Adder coherent			√	√
d) Adder Fourier flow concentrator			√	
e) Adder Quick & Painless				√
f) Adder Ubiquitous			√	

with this machine, then in principle he could construct any genome he wanted, without regard to splice points or available editing enzymes.

So he first transcribed the genome of *Mycoplasma mycoides*, editing the computer sequence with some added information (e.g. spelling his name using the one-letter peptide abbreviations for 3-basepair codons). Since the resulting 930,000 basepairs was longer than any gene machine could make (for technical reasons) he made the DNA in 1000 long strands that were "stitched together" by the enzymes in a living yeast cell. When the complete 930,000 DNA strand was finished, it was extracted from the yeast cell and transplanted into a living *M. mycoides* whose own DNA had just been removed. This is akin to changing horses in the middle of a stream, or rewiring the factory without turning off the power. Success was determined when the "synthetic" *M. mycoides* reproduced and also incorporated the "watermarked" DNA. Venter did not so much create life as modify it, and in a process that carefully added subunits together so as to not disturb the existing information. It is a case study of how to coherently add information without allowing entropic decay. Thus the flow of information through Fourier space not only concentrates information (10 years, labor, huge laboratory resulting in a 2-micron microbe) but also does it coherently, without entropically destroying the information already there. In physics terminology, the characteristics of a coherent adder are low-temperature storage where the information was frozen, interspersed with brief reactions where the entropy is tightly controlled by a highly coherent system while addition occurs.

Coherence Conclusions The information in the cell is so much greater than the computational capabilities of a universe-computer, that we are entitled to neglect the amount of information generated by it. Evolutionary search algorithms cannot operate more efficiently than a random search without possessing information themselves, and the search for a better search is less likely of success than looking for the original information. Since randomly finding or generating information cannot proceed faster than the minuscule 10^{120}/universe, the only way for evolution to provide an increasingly complex system is to add pre-existing information. The only way for the universe to contain this large amount information is through coherence–random access memory is insufficiently dense. Therefore the mechanism that adds information must add coherently without entropic decay, which requires the adder to be a coherent system as well. The net result of adding coherent information is a flow of information through Fourier space from large spatial and temporal scales to small scales. Temporally, the adder must freeze the information preventing entropic decay in between brief periods of coherent addition–which like dentistry, should be "quick and painless". And if the adder is operating on coherent, global information, then the adder must also be coherent and global, having a Fourier space component at least as large as the coherence it is adding. Table 2 summarizes the characteristics.

From considering these limitations, we propose 4 rules or properties of coherent adders: 1. they must have spatial coherence (the more global the better); 2. they must have temporal coherence (stasis without entropic decay); 3. they must perform addition with temporal coherence (quick and painless); and, 4. they must have spatial coherence at least as large as the system (ubiquitous).

5. COMETS AS CARRIERS OF COHERENT INFORMATION

Having spent the previous sections arguing that information must be coherent to account for the large amount in cells, and that evolution can only proceed if there be a coherent adder, we are now in a position to argue for the

Table 3. Problems with warm ponds

Information Limitations	Coherence Spatial	Temporal	Adder Spatial	Adder Temporal
a) Strong mixing	X			
b) High Temperatures		X		
c) No connection to more ponds			X	
d) Dilution of products			X	X
e) Continuous reactions				X
f) Rarity of ponds			X	

superiority of comets over "warm ponds" for the mechanism of evolution. The idea of a warm pond was initially due to Darwin, who no doubt thought of it as an ideal growth medium for bacteria, which are the preferred "first life" candidate of top-down theorists. The essentials of the warm pond do not change much even with modern theories, for they all insist on temperatures high enough to drive chemical reactions without enzymes, and for the availability of organic chemicals in solution to start the first cell. However there are many drawbacks to warm ponds, which can be seen by comparison to the criteria we adduced earlier.

First, the global extent and spatial coherence of warm ponds is limited because they are isolated from each other and from the universe. The information flow is only from 10's of meters down to a 2 microns, or about 7 orders of magnitude, compared to the 24 or so orders of magnitude above this spatial scale. Second, the temporal stasis and related Gibbs free energy are greatly reduced in a warm pond simply because it is warm, and because warm fluids mix so as to reduce the free energy.[38] The same features that attract theorists–the rapid chemical reactions due to heat–also cause rapid entropic decay.[39] Third, the addition of information is neither quick nor painless, since the pond possesses continuous heat and mixing. But addition needs to proceed, like Venter's swap of DNA, as rapidly as possible, and if possible, while the organism is in a condition of stasis. Warm ponds do not transfer information either quickly or without homogenizing/destroying spatial information in the pond. Finally, the warm ponds are not ubiquitous, but relatively rare throughout both the Earth and the Galaxy. Not only does it require a rocky body planet (rather than a gas-giant), but it requires the planet lie in the "Goldilocks Zone", that the climate is appropriate, that water is available, and an atmosphere can keep it in the liquid state. All these requirements mean that warm ponds will have severe limits on the amount of coherent information they can contain or add. We summarize these results in Table 3.

Warm ponds, unlike Venter's lab procedures, have: a) strong mixing due to being a fluid with temperature gradients that destroy spatial coherence; b) high thermal temperatures that encourage entropic decay more than favorable chemical reactions; c) limited if any information flow from scales larger than the pond; d) dilute their reaction products making further steps more difficult; e) Reactions are neither quick nor painless, but occur continuously; and, f) they are rare requiring water, atmosphere, temperature, gravity all within certain limits.

Having dispensed with warm ponds, let us try to work backwards towards a process that does meet the criteria for coherent addition. First, the larger the volume over which the mechanism operates, the more regions of Fourier space can be connected, the greater the flow of information from large to small. Such processes are ideally cosmic or at least of galactic extent. Second, the flows that move information from large to small scales must not permit entropic decay. Ideally they would occur at absolute zero, though from a cosmology viewpoint, operating at the temperature of the Cosmic Microwave Background (CMB) is as close to zero as can be thermodynamically obtained. Third, the addition step must be quick and painless, temporally brief without disturbing the coherence. Ideally it would occur instantaneously, and spatially by mere proximity without entropic "mixing". Finally, the entire process should be ubiquitous, ideally as dense as the information content itself. Since the adder must have some material machinery to implement this addition, that machinery also has coherence and information, which must have larger scales than the information it is adding. That is, every "packet" of information (1) that is distributed cosmically or galactically should be part of the flow (2) and the addition (3), nor should the flow (2) exhaust the supply of info (1), but should operate continuously from the Big Bang onward.

After composing a list of ideal properties of the Fourier flow that leads to increasing information density, which leads to life, we are finally in a position to compare it with a material body. Since Fourier flows are in "Fourier space", then for every Fourier space object there must correspond a spacetime carrier or material substance that is its transform or dual. Consideration of the universe suggest that the material carrier must be astronomical in order to span the widest possible volume, and among astronomical objects such as galaxies, stars, black holes, planets, comets, dust, and molecular clouds, comets are the best fit to the properties of this material carrier of the Fourier information flow.

Comets are global First, comets occur around most stars that condense from a proto-stellar nebula, and are thus available in galaxies. Their speeds (corresponding to the gravitational potential well of their natal star) can vary from a few km/s up to 100 km/s for globular clusters, which can connect widely separate regions of the universe, and are capable of providing the web-like structure of galaxies seen in the cosmological surveys, so that they span the largest volume possible of the universe. And if these comets carry a cyanobacterial payload with viruses, their Fourier scale has a lower limit of 0.1 micron, spanning about 32 orders of magnitude in size. Second, during interstellar transit, comets freeze to nearly the CMB temperature achieving nearly perfect temporal stasis. Their temporal coherence is therefore nearly perfect. Third, during collisions with other cometary debris which occur when comets are trapped in the gravitational well of a star, the spatial transfer of bacteria, spores or viruses occurs without mixing, and nearly instantaneously. This makes the addition of information as quick and painless as possible. Finally, if comets are not only remnants of stellar formation, but also primeval remnants of the Big Bang Nucleosynthesis (assuming more CNO than in current models), then they might also represent the 70% dark matter of the universe and be more ubiquitous than stars. Details are provided in previous publications.[3, 4, 40]

First, at speeds up to 30km/s in the arms of the Milky Way, comets can diffuse life throughout the galaxy by being captured by a star, decaying, infecting local comets, and being re-ejected on a new comet. Such an infection diffusion wave is calculated to travel across the Milky Way in about 1 billion years, so that since the formation of our galaxy, comets would have made twelve complete sweeps of the galaxy. It may also be possible for galaxies to infect each other through high speed comets ejected from globular clusters at about 100 km/s, though obviously, the earlier universe would be better for intergalactic infection since the galaxies were closer together. Greater precision in the calculation may establish a cutoff time, beyond which it is improbable for intergalactic infection to spread.

Second, infected comets are bioengineered to be better transporters. Cyanobacteria produce polysaccharide sheaths that blacken in UV exposure, increasing the thermal input, increasing the tensile strength of the crust and hence the production of liquid water, decreasing the loss of volatiles, increasing the power of jets that enable them to escape the stellar gravitational wells and provide an "anti-gravity" that increases their virial temperature. All of this means that the average properties of comets and galaxies are influenced by the lowly bacterium.

Third, if comets are primordial, as we suggest from a reconsideration of the theory of Big Bang Nucleosynthesis, then they may be responsible for the condensation of galaxies in the first 100 Myr of the universe. If these comets are also infected, then it may be said that galaxies are bio-engineered by cyanobacteria. Then those nucleated galaxies spin off more comets so that a phase transition marches through the universe at the speed of cometary diffusion, much as a comet undergoes a phase transition the first time it melts.[41–43] This shows how information flow is bidirectional in Fourier space, flowing from large to small and back to large, so that comets are not just a unidirectional flow, but a diffusion flow, a redistribution of information throughout all Fourier space. Thus the early universe shows a reverse flow, or an emergence, whereas late universe shows a concentrating flow. If this is an authentic observation, it would be a fingerprint for OOL.

What would that fingerprint look like? Since we are demonstrating how information diffuses through Fourier space, it would be appropriate to call it "scale-invariant" which is to say, fractal. Fractals are formed when a space-filling curve nevertheless has limitations on area or volume, such as diffusion-limited growth of snowflakes. Condensation of galaxies from molecular clouds is exactly such a process, and if comets played a part in it, we would expect to see the evidence of cometary tracks, like the tips of snowflakes that preferentially grow faster. In the left panel of Figure 1, we show a false-color simulation of galactic density in the early universe condensing

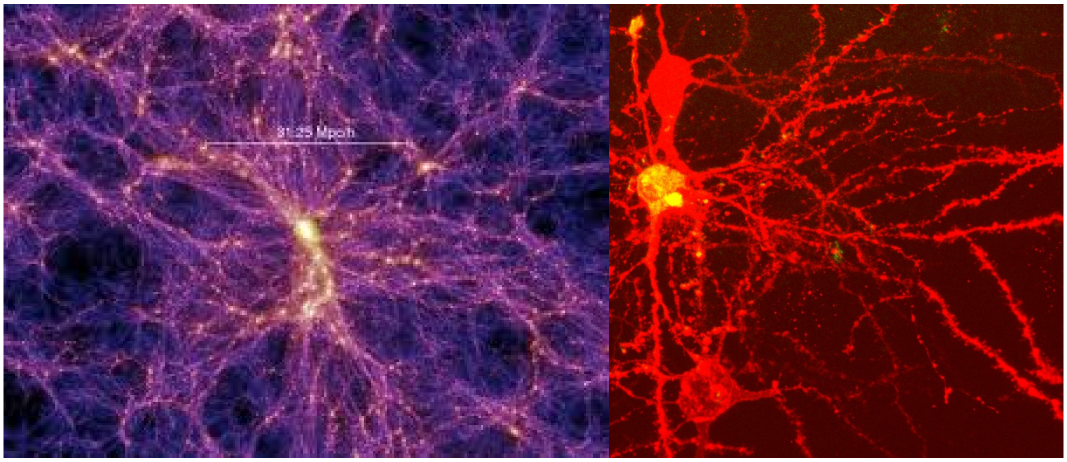

Figure 1. Left: computer simulation of the distribution of dark matter at an early point in the history of the universe. Image credit: Volker Springel/Millennium Simulation. Right: fluorescence staining of brain neurons. Image credit: neurollero/Flickr.

along lanes established by dark matter. Where several lanes intersect are greater than average density galaxy clusters. While a simulation, the results are "tuned" to agree with the observed distribution of galaxies that are now archived in public databases. The right panel shows a fluorescent stained photomicrograph of brain neurons. Both demonstrate the fractal structure of a system that must be global while delivering a limited resource and therefore collapses to a dimension more than 2 and less than 3. But rather than the dark matter in the simulation being "tuned" to achieve agreement with observations, it may be that life tunes the dark matter, achieving conditions appropriate for galaxy growth and more comets, just as a neuron grows in a way to achieve maximum information transfer and connectivity.

Comets are frozen The conclusion that comets freeze thoroughly, though not without some heat storage, is both theoretically and observationally determined. The tails of comets vanish on their way out of the solar system, someplace near the asteroid belt or Jupiter, though "outbursts" have been known to occur outside the orbit of Jupiter, depending on the time-lag and energy storage of the particular comet. Despite these semi-mysterious outbursts, most comets freeze down to a black nucleus when distant from the sun, providing a thermodynamic mechanism for putting life in the "deep freeze" until needed during a close pass of another star or sun. This behavior permits life to be held in stasis, without entropic decay, for long periods of time.

Furthermore, as we have earlier noted, comets create a concrete shell between 1/2 meter and 10's of meters thick that shields the interior of the comet from the damaging effects of cosmic rays and UV radiation during this time of stasis.[42, 43] For most of the solar comets, this deep freeze is the Oort cloud about 1 ly from the sun, and at an average temperature of just a few degrees above absolute zero. We note that the habit of comets jetting near their perihelion when they warm and melt, gives them a momentum boost that raises their aphelion even further out of the Oort cloud, and potentially ejects them from the Solar system. This effect operates somewhat like anti-gravity, assuring that comets do not spend too much time near a warm star, and keeps them frozen for long periods.

Comets melt briefly The time spent within 2AU of a G-type star is only a few months out of an orbit of 10,000 years. We can evaluate this summer season by dividing by the winter season when the comets are frozen, which give us a melt/frozen ratio of $> 10^4$. This eccentricity and hence ratio is even greater for hyperbolic orbits expected for interstellar comets at about 10^7 or 10^8. So the addition of information while a comet is melted, say, by accreting material at perihelion from previously shattered comets, which can transfer the genome of one species of cyanobacteria to another through horizontal gene transport, occurs for relatively brief moments, making the coherent information adder both quick and painless. Only for Jupiter class comets, whose aphelion has been captured by Jupiter's gravitational well and hence have orbits of about 5 years with melting periods of almost 1 year, do we have much lower numbers of about 10. These comets may not be suitable for adding, but spores and dust generated by the decay of these comets provide the raw materials that are accreted by the uninfected, Oort cloud comets.

So whatever way we view this entropic, melted phase of comets, it remains about the same length of time independent of how long the frozen journey takes. When coupled with the 1/2 meter to 10 meter crust of the comet, we find the interstellar journeys are both practical and safe for the living packages kept inside.

Comets are common We hypothesize that the destination and the journey are one, that the information transferred by a comet, and the information of the transfer itself must be nearly equivalent. Our hypothesis is that this is a consequence of the Maximum Entropy Production Principle applied to Fourier space.[44,45] For example, in modern computer codes, the program and the data are indistinguishable and of comparable size, or else it would be more efficient to move code into data or data into code. Or in a matched electrical load the external resistance must equal the internal resistance, which maximizes the power and hence the entropy production. Since the Fourier Transform that correlates spacetime includes the comets as well as the payload of the comets, therefore in some fractal dimension that mirrors the diffusion of information through Fourier space, comets must be as numerous in the cosmos as the bacteria are in the comet. Otherwise the information in transform space would not be scale-invariant, and scale-invariance is a consequence of diffusion equilibrium (of a fractional transport equation) in Fourier space. All that to say that comets should be ubiquitous and not rare, if they are to maximize their information transport.

We have observed some 10 comets in the past 300 years that have had parabolic or hyperbolic orbits, and are not gravitationally bound to the Sun as is the case of the approximately 290 elliptical orbits observed.[46] From this number, we can estimate an extra-solar "free" comet density, which when multiplied by the greater volume of interstellar space, arrives at an equivalent number of "free" comets as "bound" comets in the galaxy. So for every Earth mass of comets bound up in an Oort Cloud, there is another Earth mass of comets wandering between stars. Even though this is a big number, we think it may be dwarfed by the number of primeval comets generated in the Big Bang Nucleosynthesis of CNO which account for the 70% of non-observed baryonic matter estimated from gravitational perturbations of galaxies. This estimate is shown to be consistent with data from the collision of galaxy clusters in the Bullet galaxy cluster system.[40] In that case, interstellar comets are not one Earth mass per star, but 7 stellar masses per star, or about 10^7 more numerous and ubiquitous.

6. CONCLUSIONS

Summarizing the findings of this paper:
1) We demonstrated that OOL could be represented by information.
2) We calculated that OOL contained no less than $10^{200,000}$ "bits".
3) We reported that the universe as a random search engine could generate only 10^{120} bits.
4) We addressed this mismatch by invoking "coherence" in the information storage, represented by Fourier space information.
5) We deduced that a/biotic evolution could then be envisioned as a series of steps that added coherent information in Fourier space.
6) We adduced that adding coherent information required a coherent adder.
7) We adduced that the coherent adder could be represented by an information flow from large scale to small scale.
8) We characterized this adder as possessing 4 attributes:
a) Spatial coherence–global.
b) Temporal coherence–stasis.
c) Quick and Painless adding–short duration, spatially coherent.
d) Ubiquity–scale invariance of Fourier space.
9) We demonstrated that comets fulfill all four of the requirements above.
10) We argue that further details of the model will come from a fractal analysis of the cometary information web.

The Copernican Revolution that removed man from the center of the universe is itself overturned as we find the information flow is from the universe to us.

NOTE ADDED WHILE IN PROOF It has come to our attention that we have inconsistently applied information "bits" and information "states". Lloyd's calculation for bits of information calculable by the universe is 10^{120}, which we had incorrectly read as "number of states" and therefore incorrectly assumed 120 information bits. The number of states the universe can support is $2^{(10^{120})}$, which is sufficient to hold the information of 492,970 basepair minimal genome. Conclusion 2 should read, $10^{200,000}$ STATES can be represented with 664,000 BITS of information, and Conclusion 4 along with the section "Information Deficit" should be modified accordingly. The universe as a computer CAN compute the information in the minimal genome, and therefore the claim of non-computability and no "front-loading" must be dropped. While this makes the remaining conclusions less forceful, we stand by the conclusion of coherence.

REFERENCES

1. C. Darwin, *On the Origin of Species by Means of Natural Selection*, London: John Murray, 1859.
2. G. Wald, "The origin of life," *Scientific American* (8), pp. 44–53, 1954.
3. R. B. Sheldon and R. B. Hoover, "The cometary biosphere," in *Instruments, Methods, and Missions for Astrobiology X*, R. B. Hoover, G. V. Levin, and A. Y. Rozanov, eds., pp. 6694–0H, Proc. of SPIE Vol 6694, (Bellingham, WA), 2007.
4. R. B. Sheldon and R. B. Hoover, "Cosmological evolution: Spatial relativity and the speed of life," in *Instruments, Methods, and Missions for Astrobiology XI*, R. B. Hoover, G. V. Levin, and A. Y. Rozanov, eds., pp. 7097–41, Proc. of SPIE Vol 7097, (Bellingham, WA), 2008.
5. E. D. Schneider and D. Sagan, *Into the Cool: Energy Flow, Thermodynamics, and Life*, University of Chicago Press, Chicago, 2005.
6. A. McIntosh, "Information and entropy top-down or bottom-up development in living systems?," *Int. J. of Design & Nature and Ecodynamics* **4**(4), p. 351385, 2009.
7. N. Lane, *Life Ascending: the ten great inventions of evolution*, W. W. Norton & Co., New York, 2009.
8. C. H. Gibson, N. C. Wickramasinghe, and R. E. Schild, "First life in primordial-planet oceans: the biological big bang," *ArXiv e-prints* , Sept. 2010.
9. T. Young, *Account of the Recent Discoveries in Hieroglyphic Literature and Egyptian Antiquities*, John Murray, London, 1823.
10. D. S. Spiegel and E. L. Turner, "Life might be rare despite its early emergence on earth: a bayesian analysis of the probability of abiogenesis," *ArXiv e-prints* , July 2011.
11. D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R. Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z. Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison 3rd, H. O. Smith, and J. C. Venter, "Creation of a bacterial cell controlled by a chemically synthesized genome," *Science* **329**, pp. 52–6, Jul 02 2010.
12. S. Lloyd, "Computational capacity of the universe," *Phys. Rev. Lett.* **88**, p. 237901, May 2002.
13. D. F. Styer, "Entropy and evolution," *Am. J. Phys.* **76**, 2008.
14. R. B. Hoover, E. V. Pikuta, N. C. Wickramasinghe, M. K. Wallis, and R. B. Sheldon, "Astrobiology of comets," in *Instruments, Methods, and Missions for Astrobiology VII*, R. B. Hoover, G. V. Levin, and A. Y. Rozanov, eds., pp. 93–106, Proc. of SPIE Vol 5555, (Bellingham, WA), 2004.
15. S. J. Giovannoni, H. J. Tripp, S. Givan, M. Podar, K. L. Vergin, D. Baptista, L. Bibbs, J. Eads, T. H. Richardson, M. Noordewier, M. S. Rapp, J. M. Short, J. C. Carrington, and E. J. Mathur, "Genome streamlining in a cosmopolitan oceanic bacterium," *Science* **309**(5738), p. 12421245, 2005.
16. V. Vasas, E. Szathmry, and M. Santos, "Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life," *Proc. of Nat. Acad. Science* **107**(4), pp. 1470–1475, 2010.
17. F. Hoyle and N. C. Wickramasinghe, *Evolution from Space*, Simon & Schuster, 1981.
18. F. Hoyle and N. C. Wickramasinghe, *The Intelligent Universe*, Dent, 1983.
19. F. Hoyle and N. C. Wickramasinghe, *Astronomical Origins of Life: Steps Toward Panspermia*, Kluwer Academic Publishers, Dordrecht, 2000.
20. B. Christen, E. Abeliuk, J. M. Collier, V. S. Kalogeraki, B. Passarelli, J. A. Coller, M. J. Fero, H. H. McAdams, and L. Shapiro, "The essential genome of a bacterium," *Molecular Systems Biology* **7**(528), 2011.
21. W. A. Dembski, *No free lunch: Why specified complexity cannot be purchased without intelligence*, Rowman and Littlefield Pub. Inc., Lanham, MD, 2002.
22. W. A. Dembski and R. J. Marks, II, "The search for a search: Measuring the information cost of higher level search," *Journal of Advanced Computational Intelligence and Intelligent Informatics* **14**(5), pp. 475–486, 2010.
23. A. S. Eddington, *The Nature of the Physical World*, MacMillan, New York, 1928.
24. L. Susskind, *The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics*, Little, Brown, & Co., New York, 2008.

25. G. Montaez, W. Ewert, W. A. Dembski, and R. J. M. II, "Vivisection of the Ev computer organism: Identifying sources of active information," *Biocomplexity* **2010**(3), pp. 1–6, 2010.
26. S. Hawking and L. Mlodinow, *The Grand Design*, Bantam Books, New York, 2010.
27. J. B. Pitts, "Gauge-Invariant Localization of Infinitely Many Gravitational Energies from all Possible Auxiliary Structures, Or, Why Pseudotensors are Okay," *ArXiv e-prints* , Oct. 2009.
28. R. Lieu, "Strong-Lensing Time Delay: A New Way of Measuring Cosmic Shear," *Astrophys. J.* **674**, pp. 75–79, Feb. 2008.
29. A. A. Abdo, M. Ackermann, M. Ajello, K. Asano, W. B. Atwood, M. Axelsson, L. Baldini, J. Ballet, G. Barbiellini, M. G. Baring, and et al., "A limit on the variation of the speed of light arising from quantum gravity effects," *Nature* **462**, pp. 331–334, Nov. 2009.
30. L. Susskind, *The anthropic landscape of string theory*, pp. 247–+. Cambridge University Press, 2007.
31. V. G. Gurzadyan and R. Penrose, "Concentric circles in WMAP data may provide evidence of violent pre-Big-Bang activity," *ArXiv e-prints* , Nov. 2010.
32. F. Hoyle, G. Burbidge, and J. V. Narlikar, *A different approach to cosmology : from a static universe through the big bang towards reality*, Cambridge Univ. Press, London, 2000.
33. R. Jastrow, *God and the Astronomers*, Norton, New York, 1978.
34. R. Penrose, *The emperor's new mind. Concerning computers, minds and laws of physics*, University Press, Oxford, 1989.
35. G. Amelino-Camelia, L. Freidel, J. Kowalski-Glikman, and L. Smolin, "Relative locality: a deepening of the relativity principle," *General Relativity and Gravitation* , pp. 125–+, June 2011.
36. C. E. Shannon, "Prediction and entropy in written english," in *Claude Shannon collected papers*, W. A. D. Sloane N. J. A., ed., p. 583, IEEE Press, Piscataway, NJ, 1993.
37. L. Margulis, *The symbiotic planet*, Basic Books, New York, 1998.
38. J. Attwater, A. Wochner, V. B. Pinheiro, A. Coulson, and P. Holliger, "Ice as a protocellular medium for rna replication," *Nature Communications* , 2010.
39. J. Sanford, *Genetic Entropy and the Mystery of the Genome*, Feed My Sheep, Inc, 2008.
40. R. B. Sheldon, "Comets, information, and the origin of life," in *Origins: Genesis, Evolution and Diversity of Life*, J. Seckbach and R. Gordon, eds., *COLE*, Springer, 2011.
41. R. B. Hoover, E. V. Pikuta, N. C. Wickramasinghe, M. K. Wallis, and R. B. Sheldon, "Astrobiology of comets," in *Instruments, Methods and Missions for Astrobiology VIII*, R. Hoover, G. Levin, and A. Rozanov, eds., **SPIE 5555**, SPIE, Bellingham WA, 2004.
42. R. B. Sheldon and R. B. Hoover, "Evidence for liquid water on comets," in *Instruments, Methods, and Missions for Astrobiology VIII*, R. B. Hoover, G. V. Levin, and A. Y. Rozanov, eds., pp. 196–206, Proc. of SPIE Vol 5906A, (Bellingham, WA), 2005.
43. R. B. Sheldon and R. B. Hoover, "Implications of cometary water: Deep impact, stardust and hayabusa," in *Instruments, Methods, and Missions for Astrobiology IX*, R. B. Hoover, G. V. Levin, and A. Y. Rozanov, eds., pp. 6309–0L, Proc. of SPIE Vol 6309, (Bellingham, WA), 2006.
44. H. Ozawa, A. Ohmura, R. D. Lorenz, and T. Pujol, "The second law of thermodynamics and the global climate system: A review of the maximum entropy production principle," *Rev. Geophys.* **41**(4), p. 1018, 2003.
45. L. Martyushev and V. Seleznev, "Maximum entropy production principle in physics, chemistry and biology," *Physics Reports* **426**(1), pp. 1–45, 2006.
46. L. L. Wilkening, *Comets*, Univ. Arizona Press, 1982.