An Optimal, Universal and Agnostic Decoding Method for Message Reconstruction, Bio and Technosignature Detection

Hector Zenil*[1,2,3,4,5], Alyssa Adams[5,6], Felipe S. Abrahão[3,7,8]

[1] School of Biomedical Engineering and Imaging Sciences, King's College London, U.K.
[2] The Alan Turing Institute, British Library, U.K.
[3] Oxford Immune Algorithmics, U.K.
[4] Algorithmic Dynamics Lab, Karolinska Institutet, Sweden.
[5] Cross Labs, Cross Compass, Kyoto, Japan.
[6] John W. and Jeanne M. Rowe Center for Research in Virology, Morgridge Institute for Research, University of Wisconsin–Madison, U.S.
[7] Centre for Logic, Epistemology and the History of Science, University of Campinas, Brazil.
[8] DEXL, National Laboratory for Scientific Computing, Brazil.

Abstract

We present a signal reconstruction method for zero-knowledge one-way communication channels in which a receiver aims to interpret a message sent by an unknown source about which no prior knowledge is available and to which no return message can be sent. Our reconstruction method is agnostic vis-à-vis the arbitrarily chosen encoding-decoding scheme and other observer-dependent characteristics, such as the arbitrarily chosen computation model or underlying mathematical theory. We investigate how non-random messages may encode information about the physical properties, such as dimension and length scales of the space in which a signal or message may have been originally encoded, embedded, or generated. We argue that our results have applications to life and technosignature detection and to coding theory in general.

Keywords: algorithmic information dynamics, causal deconvolution, zero-knowledge communication, technosignatures, signal processing, perturbation analysis, universal distribution, intelligent signal detection, biosignatures

---

*Corresponding author. Email: hector.zenil@cs.ox.ac.uk

1 Introduction

For the past 50 years, astronomers have been sweeping the skies using radio telescopes in hopes of stumbling across a message from an alien civilisation. More recently, space missions to collect samples and data from Mars, Titan, Europa, and others have been focused on detecting possible biosignatures [1–5]. The James Webb Space Telescope has been outfitted with sensors and measurement devices to scout potentially habitable solar systems [6, 7]. For the first time in human history, the detection of extraterrestrial life is at the forefront of space exploration. In the novel Contact, written by Carl Sagan and later made into a film, an extraterrestrial signal is received that is encoded in three dimensions. A good deal of the story has to do with how scientists serendipitously deduce that the signal encodes an object in three dimensions after spending months or years trying to fit it in two dimensions.

Here we show that the more a message is removed from randomness the easier is to derive its native geometric and topological dimensions, and that such an approach derives its fruitful properties from the universality and limitations of algorithmic compression algorithms. These properties include being agnostic (in the asymptotic limit when computational resources are unbounded) vis-á-vis: the encoding-decoding scheme used to encode the message by the emitter agent; computation model and the programming language chosen in order to implement our proposed method; the computable (or semi-computable) method of approximation to algorithmic complexity; any arbitrarily chosen (computable) probability measure of the events; and the formal theory able to formalize the mathematical statements. While the arbitrary choice of each of these constitutes an observer dependency, this dependency is cancelled out in the long run without the need of resorting to special observers [8].

All of our knowledge about life is based on what occurs here on Earth [9–17]. Some researchers consider technosignature detection a more promising method of detecting life than other biosignatures because of their longevity [18–20]. However, some researchers worry that our current theoretical toolkit for data analysis may not be sophisticated enough [21–23]. Most of the discussion has been centred around the technical justification of hardware choices and the technicalities of detection (e.g. the frequency) and not the nature (i.e., original structural characteristics or underlying mathematical properties) of the signal itself. Thus, for example, the search for extraterrestrial signals has been mostly focused on narrow-band signals (mostly pulses) a few Hertz-wide (or narrower). Natural cosmic noisemakers, such as pulsars, quasars, and the turbulent, thin interstellar gas of our own Milky Way do not make radio signals that are this narrow, which is a reasonable justification of the bounds of the physical signal search. On Earth, we also find local signal-noise makers such as animals, organs, neurons, and cells. However, beyond identifying very basic statistical patterns, not much progress has been made on the qualitative semantic aspects of signals and communication, let alone on building a universal framework within which the question can be explored and analysed.

Each different species on Earth has its own different set of sensors that it uses to interact with its environment. To put this in terms of signals or messages between senders and receivers, different species emit signals based on the mechanisms that allow them to do so [24, 25]. They also decode messages based on their sensory capabilities, and if each species has a unique set of senses, then we can say that each species decodes the same message differently. The 52Hz whale's messages are decoded differently between humans monitoring the oceans and other whales (who do not sense the messages at all).

One can say that that if a signal is intentional, it might be decipherable. In order to intentionally send or receive a signal over interstellar distances, it is reasonable to assume a civilisation must understand basic science and mathematics, or at least must evolve in such that these are somehow encoded into a message. Hence, a message from another civilisation might use some similar framework to science and mathematics to build up a common ground with respect to other societies. Signals sent by a civilisation for its own purposes may be impossible for us to unravel but our own mathematical tools may help.

In this paper, we introduce a method based upon the principles of (algorithmic) information theory that is aimed at sweeping over various possible encoding and decoding schemes to test our current limits on signal interpretation as we attempt to reconstruct the original partition (i.e., multidimensional space) in which the original message was given its "meaning" by the emitter agent. This *semantic* characteristic refers to the (algorithmic) information about the context or real-world correspondents of the emitter that the signal sent by this emitter is trying to convey to the receiver. The real-world correspondents that we particularly investigate are the objects grounded in their respective contexts, that is, embedded into their respective multidimensional spaces as the emitter originally intended. See the Sup. Inf. for a formal introduction to these concepts.

One-way communication channels are those for which the receiver cannot (in principle or in practice) send any signal back to the emitter in order to help or facilitate the decoding process of the first message sent by the emitter. We demonstrate how signals and messages may be reconstructed by deriving the number of dimensions and the scale of each length of an object from examples ranging from text to images embedded in multiple dimensions, showing a connection between irreducible information content, syntax, geometry, semantics, and topology. In addition, this article also shows results that are not only agnostic vis-á-vis prior knowledge of encoding-decoding schemes, but also demonstrate sufficient conditions in this zero-knowledge scenario for enabling the reconstruction of the original *message* (i.e., the original object embedded into the original multidimensional space to be conveyed to the receiver) in one-way communication channels.

Zero-knowledge communication occurs when the receiver agent is able to correctly interpret the received signal as the originally intended message sent by the emitter agent, given that the receiver has no knowledge about the encoding-decoding scheme chosen by the emitter. Notice that this condition of no prior information about the emitter agent only applies to encoding-decoding schemes, and consequentially also to the original multidimensional space and object that are unknown to the receiver before any communication takes place. However, as discussed in the Sup. Inf., this zero-knowledge condition does not mean other assumptions regarding the unknown emitter agent are not being considered by the receiver agent. Also, the reader should not confuse zero-knowledge communication (ZKC) with zero-knowledge proof (ZKP), which is commonly studied in cryptography [26–28]. Actually, one may consider ZKP and ZKC as kindred mathematical problems but as diametrically opposing counterparts with regard to the acquisition of knowledge.

The method is based on the principles of Algorithmic Information Dynamics (AID) [8, 29, 30], and it consists of a perturbation analysis of a received signal. AID is based upon the principles of information theory and the mechanisms of algorithmic probability and the universal distribution, a formal approach to a type of Artificial General/Super Intelligence that requires the massive production of a universal distribution, the mother of all models [31], to build a very large (semi-computable) model of computable approximate models. The underlying idea is that a computable model is a causal explanation of a piece of data. As a result of each perturbation applied to the received signal, a new computable approximate model is built, and then compared against the observation [30, 32]. This process then builds a large landscape of possible candidate models (along with their respective approximate complexity values) from which one can infer the best model. In the particular context investigated in this article, the partition (i.e., multidimensional space) for which the message displays the lowest complexity indicates the original partition that the received signal stream encodes.

Overarching frameworks such as the one introduced here can be useful not only for signal detection but also for signal deconvolution from local information in, e.g., biology (what chemical signals among cells mean). In previous work, we showed how the same technology can be used to disentangle the 3D structure of DNA and genomic information [33].

2 Perturbation analysis of compressed data

To showcase the relationship between compression, algorithmic complexity, and encoding methods, introducing the method put forward in this work, we took real sentences from Darwin's "Origin of Species" and binarised them in two different ways according to arbitrary ASCII-to-binary functions. In one encoding (left column in Fig. A1), ASCII vowels are assigned a 1 and all others 0. For the other encoding (right column), spaces are assigned a 1 and the rest 0. The compression, block entropy, and block decomposition method (BDM) are measured for each string. Then, for each string, each value is perturbed (1 to a 0, or vice-versa) one at a time. The resulting change in values from the original binary sequence to the perturbed ones are shown as different coloured lines for each panel.

The results in Fig. A1 demonstrate that most single-bit perturbations in a sequence that encodes a message result in a more random-looking sequence. These results vary depending on what mapping we choose to binarise the sequences, but in general, any perturbation generally increases randomness across various mappings (we also explored other arbitrary binarisation functions, with similar results). In other words, this example shows that sequences that carry meaningful content are quantitatively different according to various measures. Any deviation (i.e., perturbation) from the encoded sequences causes the randomness to increase, which suggests that an incoming signal with low randomness is more likely to encode a message than one that has more randomness.

Fig. A2 demonstrates the same phenomenon shown in non-random text but in audio signals, the results using an audio recording of the words spoken from Apollo 13, "Houston, we've had a problem" transmitted on April 13, 1970 at a sample rate frequency of 11025 Hz on a single channel. When compressed, the message is 106 250 bytes and 106 320 characters long. Here, the message was scrambled several times. The histograms on the bottom left show the different compression lengths in bytes and string lengths of the scrambled messages. Each of these scramblings results in an increase in randomness. Additionally, if only the beginning and end of each word are perturbed, the resulting randomness increases smoothly with each perturbation, as shown in the bottom right plot. These results indicate that perturbations of the original message, including scrambling elements of the original message in a different order, will result in a message of higher complexity and thus greater randomness. As a result, the original encoding of the message (the one with an interpretable meaning) is the one with the lowest complexity.

3 On the information content of the Arecibo message

In 1974, the bitmap image in Fig. A3 was sent into space as a radio signal from the Arecibo radio telescope. At the left-hand end of the image is a version of the pattern of digits from page 117– but it is distorted so it has no obvious nested structure. In the image, there are atomic numbers for various elements and bitvectors for components of DNA. Under these, there are rough pictorial representations of a DNA molecule, a human, and the telescope. All these images seem to depend almost completely on human visual encoding/decoding conventions. Without any sort of human context, including any indication that these are pictorial representations based on human vision, their meaning would be essentially impossible to recognise. This is especially true for message receivers who may not possess visual recognition capabilities, at least not visual capabilities that are similar to our own vision.

In addition, this message was reformatted from its original bidimensional state and sent as a string of 1,679 binary digits. The reasoning was that an alien civilisation with a mathematical system that receives the message will recognise 1,679 as a semiprime number–a multiple of 23 and 73. Even if the original message could not be reconstructed in its original and intended encoding, the receiver would recognise the message as having some mathematical significance that is unlikely to be a random signal from outer space.

Fig. A4 shows how the method provides a robust approach to identifying the best way to decode a signal back to its original multidimensional space, by sweeping over a large variety of possible dimensional configurations and measuring the resulting information content and algorithmic complexity of the candidate messages by taking the lowest algorithmic complexity configurations under the assumption that the original message is algorithmically not random (i.e., that the original message is compressible). As shown in [34] for graphs, this is because if the original message is random, then the likelihood that a rearrangement (i.e., a new partition) will lead to a low complexity configuration is very low. On the other hand, if the original message is sufficiently compressible, and therefore not random, most partitions will lead to configurations of greater algorithmic randomness.

Fig. A5 shows the method's noise resistance to indicate the candidate partition and the precise lengths of the original message. Successful amplification of the signal is shown in Fig. A6, with BDM outperforming Compression and Shannon Entropy in the face of additive noise, with Compression showing insensitivity to the original signal at about 10% of bits flipped, and Shannon Entropy diminishing faster than BDM but slower than Compression. In all cases, different levels of robustness at deconvolving the image dimensions are shown.

Investigating this further, Fig. A7 shows six bidimensional images along with their original numerical dimension. The values of the three complexity measures over possible partitions are shown below the original numerical dimension for ease of comparison. The method is invariant with regard to linear transformations such as encoding, as shown in images 2 and 4 of Fig. A7. Drops in complexity (downward spikes) indicate candidate dimensions for the original encoding dimensions, and thus the decoding dimension that would process the signal such that it produces the lowest-complexity message. BDM outperforms compression and entropy and correctly identifies the original message (image), encoding for 4 out of 6 images.

Some of these image decodings for the space shuttle image are shown in Fig. A8. Each panel shows random realignments of the correct dimensions along with the original image (i.e., object) embedded into these new partitions, with their respective BDM values displayed above each image. When compared to the other bidimensional spaces, the one that is most similar to the original message will have a lower complexity value.

By translating (encoding) a colour image into binary black and white pixels, this method picks an image decoding that is similar enough to the original image (at least similar enough visually). Fig. A9 shows the original image on the left, the results of the method on the right, and the resulting image selected from the spike in BDM on the bottom. A mirror-like image is selected as the first candidate, followed by the correct one. This kind of spiking also suggests the original partition can be inferred via smaller, non-random spikes at multiples or divisors of the native partition (500). In this case, small spikes at half the original partition (250) provide clues to the original embedding in 2D.

4 Discussion

This work advances a practical and theoretical framework that relates information, entropy, complexity, and semantics that can be extended and hence put to multiple uses in signal deconvolution, bio- and technosignature detection, cryptography, and coding theory. Our methods show how the receiver can decode the (multidimensional) space into which the original message was sent via zero-knowledge one-way communication channels.

We believe that the present work is only one example of how our mathematical and computational framework can be used in all areas of inverse problems as an approach to a large universal generative model able to instantiate Artificial General Intelligence from first principles, particularly in application to some specific cases of message reconstruction in which prior knowledge about the source is very limited. Such results relate information theory to fundamental areas of mathematics, such as geometry and topology, by means of compression and algorithmic probability.

Acknowledgements. Felipe S. Abrahão acknowledges support from the São Paulo Research Foundation (FAPESP), grants 2021/14501-8 and 2023/05593-1.

Author Contributions:. HZ conceived the theory, methods and experiments; HZ performed most of the experiments, with support from AA. HZ and FA developed the theoretical framework.

References

[1] Uthamacumaran, A., S. Abrahão, F., A. Kiani, N. & Zenil, H. On the salient limitations of the methods of assembly theory and their classification of molecular biosignatures. *arXiv Preprints* **arXiv:2210.00901 [cs.IT]** (2022). URL https://doi.org/10.48550/arXiv.2210.00901.

[2] Kiang, N. Photosynthesis and astrobiology: Looking for life elsewhere **36**, 24–30 (2014).

[3] Meadows, V. S. *et al.* The Search for Habitable Environments and Life in the Universe **2010**, 201 (2009). URL https://ui.adsabs.harvard.edu/abs/2009astro2010S.201M.

[4] National Academies of Sciences, Engineering, and Medicine, Division on Engineering and Physical Sciences, Space Studies Board & Committee on Astrobiology Science Strategy for the Search for Life in the Universe. *An Astrobiology Strategy for the Search for Life in the Universe* (National Academies Press (US), 2018). URL http://www.ncbi.nlm.nih.gov/books/NBK540091/. 30986006.

[5] Rojo, S. Center for life detection (cld) – research and service (2022). URL http://www.nasa.gov/ames/cld.

[6] Friday, C. H. . P., 22, A. & 2022. How the James Webb Space Telescope will search for extraterrestrial l. URL https://astronomy.com/news/2022/04/how-the-james-webb-space-telescope-will-search-for-extraterrestrial-life.

[7] Haqq-Misra, J. *et al.* Detectability of Chlorofluorocarbons in the Atmospheres of Habitable M-dwarf Planets **3**, 60 (2022).

[8] Abrahão, F. S. & Zenil, H. Emergence and algorithmic information dynamics of systems and observers. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **380** (2022).

[9] Zenil, H., Gershenson, C., Marshall, J. & Rosenblueth, D. Life as thermodynamic evidence of algorithmic structure in natural environments **14** (2012).

[10] Bedau, M. A. *et al.* Open problems in artificial life **6**, 363–376 (2000).

[11] Cleland, C. E. & Chyba, C. F. Defining life **32**, 387–393 (2002).

[12] Dupré, J. *The Metaphysics of Biology* (Cambridge University Press, 2021).

[13] Kempes, C. P. & Krakauer, D. C. The Multiple Paths to Multiple Life **89**, 415–426 (2021).

[14] Mariscal, C. *et al.* Hidden Concepts in the History and Philosophy of Origins-of-Life Studies: A Workshop Report **49**, 111–145 (2019).

[15] Ruiz-Mirazo, K., Peretó, J. & Moreno, A. A universal definition of life: Autonomy and open-ended evolution **34**, 323–346 (2004).

[16] Walker, S. I. & Davies, P. C. The algorithmic origins of life **10**, 20120869 (2013).

[17] Witzany, G. What is Life? **Frontiers in Astronomy and Space Sciences, 7** (2020).

[18] Zenil, H., Delahaye, J.-P. & Gaucherel, C. Image information content characterization and classification by physical complexity **17**, 26–42 (2013).

[19] Haqq-Misra, J. *et al.* Searching for technosignatures in exoplanetary systems with current and future missions **198**, 194–207 (2022).

[20] Wright, J. T. *et al.* The Case for Technosignatures: Why They May Be Abundant, Long-lived, Highly Detectable, and Unambiguous **927**, L30 (2022).

[21] Botta, O. *et al.* (eds) *Strategies of Life Detection* Vol. 25 of *Space Sciences Series of ISSI* (Springer US, 2008).

[22] Enya, K., Yamagishi, A., Kobayashi, K. & Yoshimura, Y. Comparative study of methods for detecting extraterrestrial life in exploration mission of Mars and the solar system **34**, 53–67 (2022).

[23] Neveu, M., Hays, L. E., Voytek, M. A., New, M. H. & Schulte, M. D. The Ladder of Life Detection **18**, 1375–1402 (2018).

[24] Nagel, T. What Is It Like to Be a Bat? **83**, 435–450 (1974).

[25] Yong, E. *An Immense World: How Animal Senses Reveal the Hidden Realms around Us* First edition edn (Random House, 2022).

[26] Buchanan, W. *Cryptography* 1 edn (River Publishers, New York, 2022).

[27] Salil Vadhan. *A study of statistical zero-knowledge proofs* (Springer, Berlin, 2023). URL https://link.springer.com/book/9783540713739.

[28] Allender, E., Hirahara, S. & Tirumala, H. Tauman Kalai, Y. (ed.) *Kolmogorov Complexity Characterizes Statistical Zero Knowledge.* (ed.Tauman Kalai, Y.) *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, Vol. 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 3:1–3:19 (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2023).

[29] Zenil, H., Kiani, N. A., Abrahão, F. S. & Tegnér, J. N. Algorithmic Information Dynamics. *Scholarpedia* **15**, 53143 (2020).

[30] Zenil, H., Kiani, N. A., Zea, A. A. & Tegnér, J. Causal deconvolution by algorithmic generative models. *Nature Machine Intelligence* **1**, 58–66 (2019).

[31] Kirchherr, W., Li, M. & Vitányi, P. The miraculous universal distribution **19**, 7–15 (1997).

[32] Zenil, H. *et al.* An Algorithmic Information Calculus for Causal Discovery and Reprogramming Systems **19** (2019).

[33] Zenil, H. & Minary, P. Training-free measures based on algorithmic probability identify high nucleosome occupancy in dna sequences **gkz750** (2019).

[34] Zenil, H., Kiani, N. A. & Tegnér, J. The Thermodynamics of Network Coding, and an Algorithmic Refinement of the Principle of Maximum Entropy. *Entropy* **21**, 560 (2019).

[35] Zenil, H. *et al.* Minimal Algorithmic Information Loss Methods for Dimension Reduction, Feature Selection and Network Sparsification. *arXiv Preprints* **arXiv:1802.05843 [cs.DS]** (2022). URL https://doi.org/10.48550/arXiv.1802.05843.

Appendix A Figures

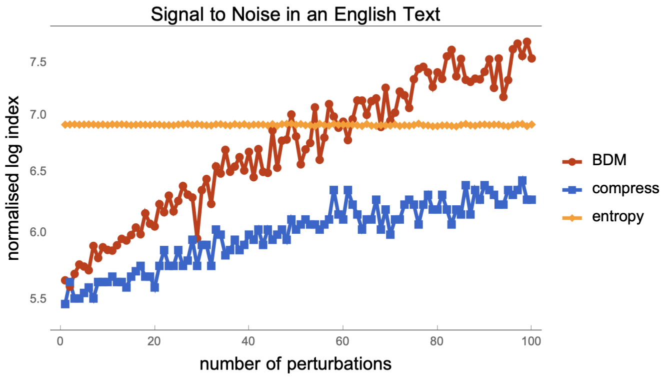


Fig. A1: Algorithmic perturbation analysis on short sentences of only 400 characters from Darwin's "Origin of Species" are converted from ASCII into their corresponding binary values. Changing a single character with probability $1/n$ for growing n from $n = 1$ (original test, axis $x = 0$) to $n = 4$ (a quarter of the text is mapped to a random character from the same vocabulary of the original text, axis $x = 100$ letters changed or perturbed) causes the resulting size of the compressed sequence index to increase, with BDM the most sensitive and Shannon entropy unable to capture any difference. This is because signals carrying meaning are far removed from randomness, and random perturbations make the text more random [35] even when the methods know nothing about words, grammar or anything linguistic.

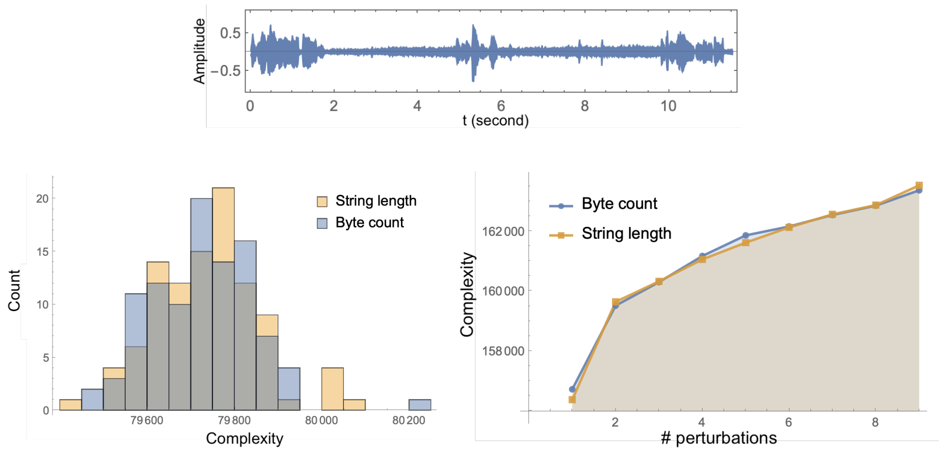


Fig. A2: *Top:* Sound waves of the words spoken from Apollo 13, "Houston, we've had a problem." transmitted on April 13, 1970 at a sample rate frequency of 11025 Hz on a single channel. *Bottom left:* Histograms of the change in compression lengths of scrambled versions of the same message from the length of the original compressed message. *Bottom right:* Small perturbations of the original message by scrambling only the beginning and end of each word shows a smooth transition from low to high randomness. The lowest complexity signal indicates the correct (original) signal. The file was processed in FLAC format (Free Lossless Audio Codec) from a lossless file, with no audio data discarded.

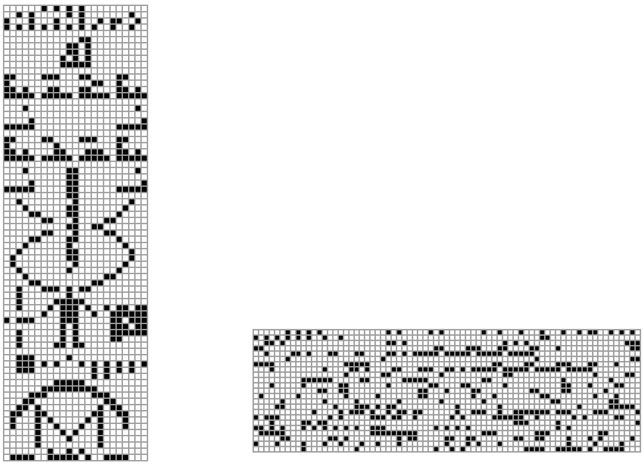


Fig. A3: *Left:* The original Arecibo message intended to be reconstructed, but sent as a linear stream from the radiotelescope in Arecibo, Puerto Rico. The 1,679 bits are meant to be arranged into 23 columns of 73 rows, 23 and 73 being two prime numbers which when multiplied together equal 1,679. *Right:* If the stream is instead arranged into 23 rows and 73 columns, the original visual interpretations of the message are scrambled, which may result in a figure that is closer to being statistically random. What we show is that the message is still there, concealed, and can be deciphered by algorithmic deconvolution.

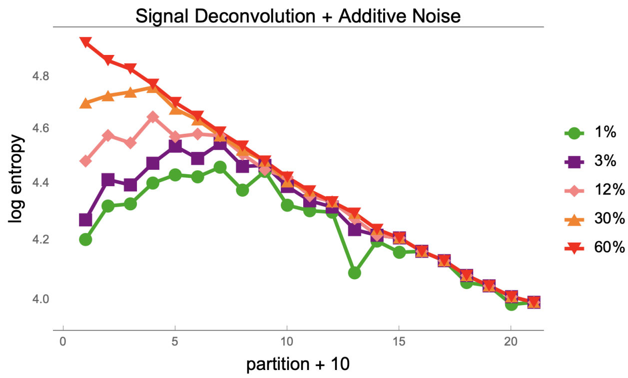


Fig. A5: The method's resilience in the face of some noise. At 3% of the bits of the original $23 \times 73 = 1679$-pixel image randomly flipped (which means about 1.5% were binary negated), the method remains sensitive and displays a small downward spike at the 23 value, uncovering its length, but the signal gets lost when more bits are flipped.

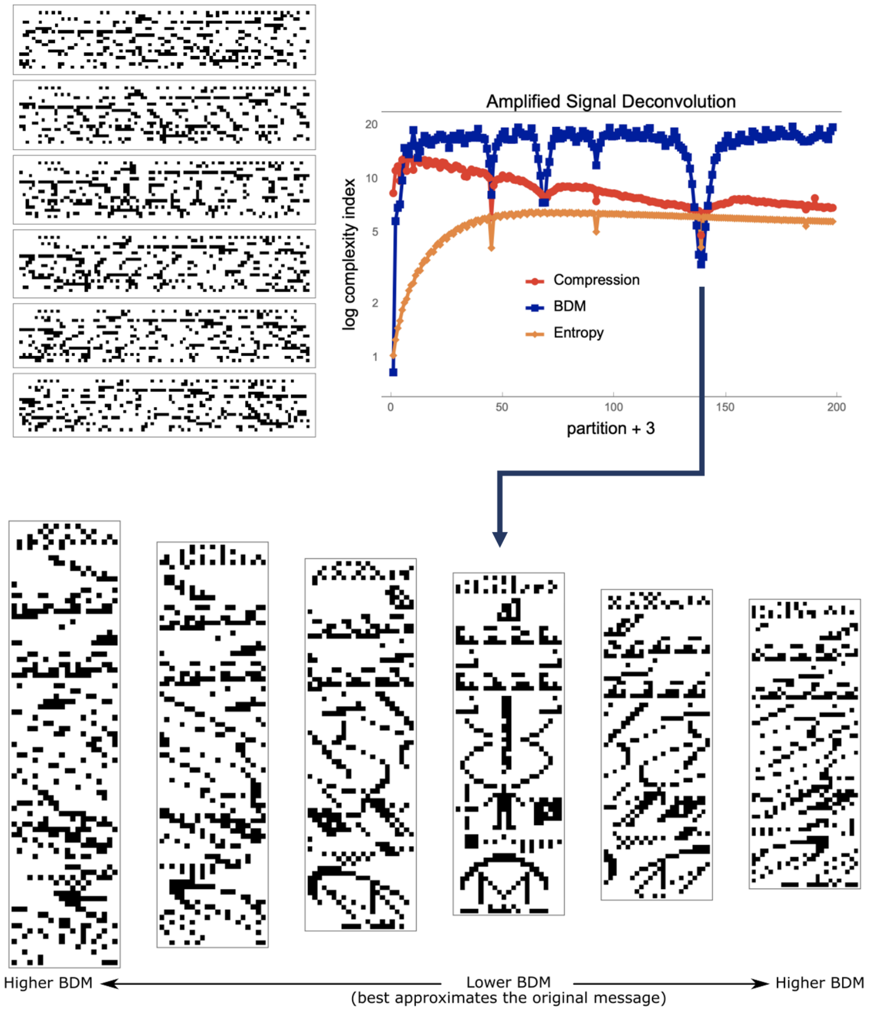
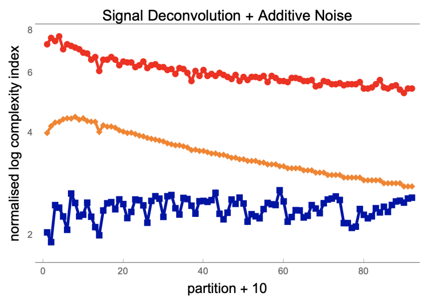


Fig. A4: *Top left:* Most possible partitions result in random-appearing configurations with high corresponding complexity, indicating measurable randomness. *Bottom:* Some partitions will approximate the originally encoded meaning (third from the right). Other configurations result in images with higher complexity values. This sequence of images shows the images in the approximate vicinity of the correct bidimensional configuration (i.e., partition) and illustrates fast convergence to low complexity. *Top right:* By using different information indexes across different configurations, a downward-pointing spike will indicate message (image) configurations that correspond to low-complexity image(s). This allows a prior-knowledge-agnostic and objective method to infer a message's original encoding. Of the various measures, BDM, combining classical information (entropy) for long ranges and a measure motivated by algorithmic probability for short ranges, is the most sensitive and accurate in this regard. Traditional compression and entropy also indicate the right configuration amongst the top spiking candidates. The ratio of noise-to-signal was amplified in favour of the hidden structure by multiplying the original image size by 6 for both length and height.



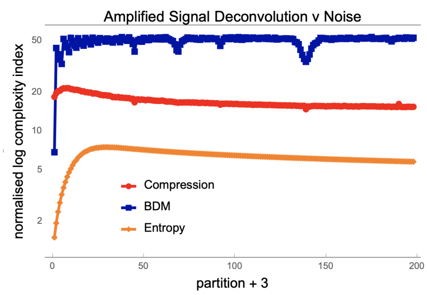


Fig. A6: Applying 3 different quantitative methods, it is shown (left) that they are highly insensitive to signal and highly sensitive to noise (16.5% of pixels randomly flipped). However, by growing the original image (right) by a factor of 6 on each dimension (i.e. 1 pixel becomes 6×6), the methods are less sensitive to noise and more sensitive to signal, with BDM significantly outperforming Compression, and Shannon Entropy showing sensitivity at up to 60% pixels flipped (hence about 30% of the original image) versus Compression and Shannon Entropy, that are about 50% sensitive. Downward spikes (right) are shown at $23 \times 6 = 141$.

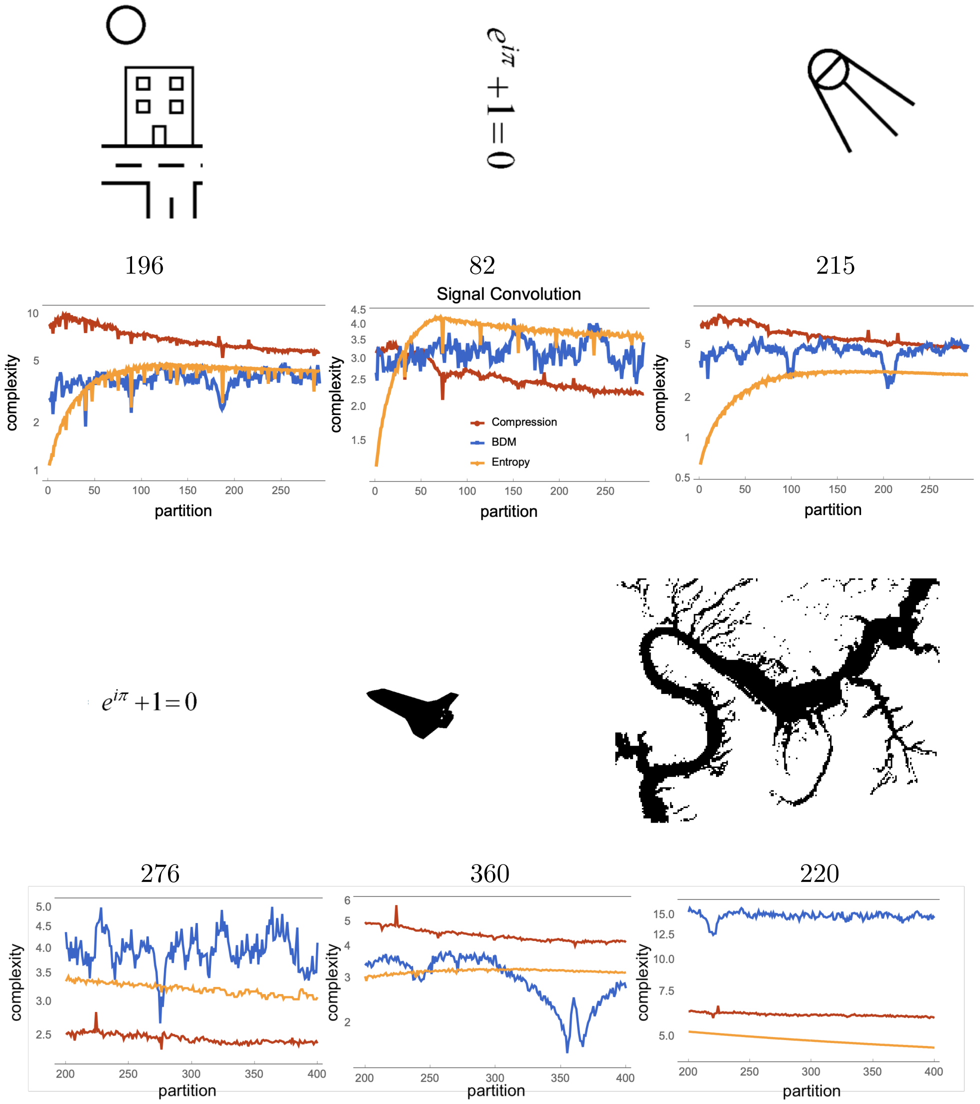


Fig. A7: Six 2D images (labelled 1 through 6 going from left to right, top to bottom) of very different nature, including a demonstration of linear-transformation invariance conforming to the underlying theory (in this case, rotation of a mathematical formula). Size invariance has actually shown amplification of signal-to-noise difference. Under each image is its correct numerical (first) dimension. The values of the three complexity indexes over possible partitions are shown below the original numerical dimension for ease of comparison. Downward spikes indicate candidates for possible original partitions. In all cases, the correct dimension value is among the top three candidates, with BDM outperforming in 4 out of 6 cases at indicating the top candidate.

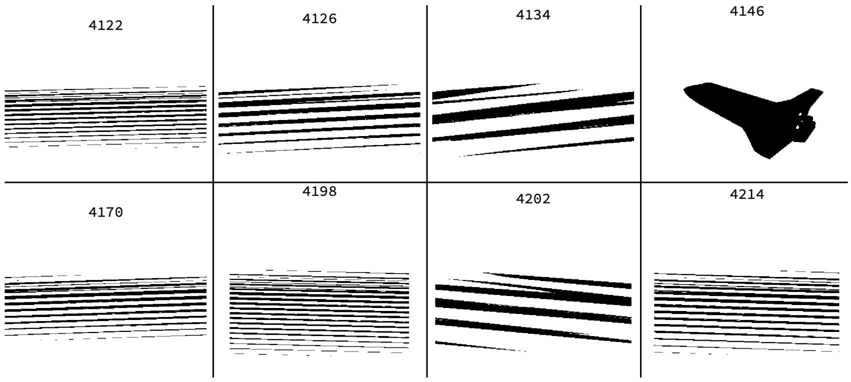


Fig. A8: Sequence of lowest complexity (BDM value on top in bits) partitions of a black and white image of a space shuttle, with the top right partition as the correct reconstruction and other candidates cases in which an alignment occurs at some multiple of the correct dimension.

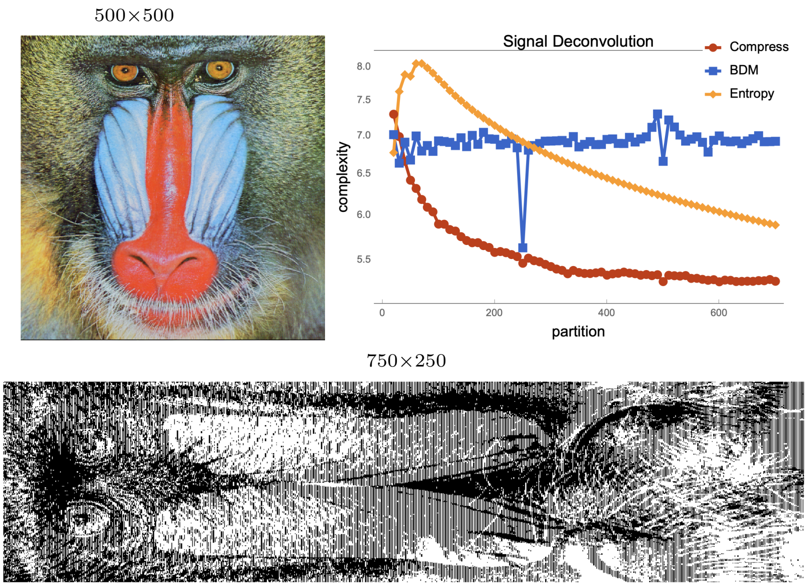


Fig. A9: *Top Left:* An original 2D image message and *top right:* the complexity results of its various image reconstructions via linear signal decomposition (in binary, similar to how the Arecibo message was sent) into distinct partitions. BDM spikes prominently at the correct right partition with bidimensional configuration of 500 × 500 pixels, contrasting with a mirror-like image at 250 × 750 pixels (*bottom*, rotated 45 degrees counterclockwise).