# An alternative/complementary approach to finding the GUT symmetry group of physics

**Kingsley J.A. Cox**

**Independent Researcher, Stony Brook, NY**

**email: kcox@syndar.org**

**Introduction**

The Universe and the brain have at least one thing in common – they both create the Universe, the actual Universe creates itself, and the brain creates a version of the Universe. With this in mind, is there anything that the brain can say about the Universe? If the cortex is a causal engine, then it may be the right kind of neural network — one that removes symmetries and seeks generative structure - that we need to discover **t**he hidden causes of complex systems such as particle interactions.

Symmetries of the universe play a big role in physics and this is seen by the current description of the Standard Model as the symmetry group SU(3)xSU(2)xU(1) (Kibble (2014)). A further extension of this group to a larger one (what is termed the GUT symmetry group) has been put forward as a possible way to explain the zoo of fundamental particles including the graviton, a putative force carrier of the gravitational force (Croon et al (2019)).

Here we outline a brief possible method of assisting the search for the symmetry group underlying physics. It uses the observation that the brain, a multilayer Neural Network (NN), can be thought of as a symmetry reduction device (Laird et al (2025), Yang et al (2024)). The brain gets exposed to input data and removes symmetries contained in the input dataset layer by layer ultimately uncovering 'causes' of the data (e.g. objects). If we could examine the connections in each layer of the brain we could in principle work backwards from the 'causes' to the original symmetry group that describes all the symmetries present in the input dataset – the 'Root' symmetry group of the brain. The problem is that the brain is not exposed to all the data that the universe has to offer and so its Root group will not be the symmetry group of the universe. Furthermore, it would be very difficult to analyse the connections in a biological system like the brain. However, a sufficiently large NN allowed to learn on a much larger set of data would be amenable to analysis (since it is digital), and it may be possible uncover its Root symmetry group. With a sufficiently large diverse set of inputs, this Root group may be close to, or even the same as, the actual GUT group of the universe.

How the brain is able to make sense of its environment has been extensively studied both experimentally and theoretically over the past few decades. One of the most interesting findings is that the neocortex is arranged in layers in a feed-forward manner with information being processed at each layer and then passed onto the next layer. At each layer some information is extracted by each neuron in the layer and that information is distributed to other neurons in higher layers and the same layer. There are also extensive feedback pathways but their exact role is as yet unclear. The main engine for this process is thought to be 'Hebbian Learning' where a connection between two neurons changes strength depending on how closely the pre and post neurons fire together. This very simple rule underlies most of the neural network AI models and can be shown to extract information from data and solve complex problems like image recognition.

**Natural scenes**

Hypothesis: A multi-layer NN that recognizes objects goes through a process at each layer of removing sets of symmetries (figure 1).

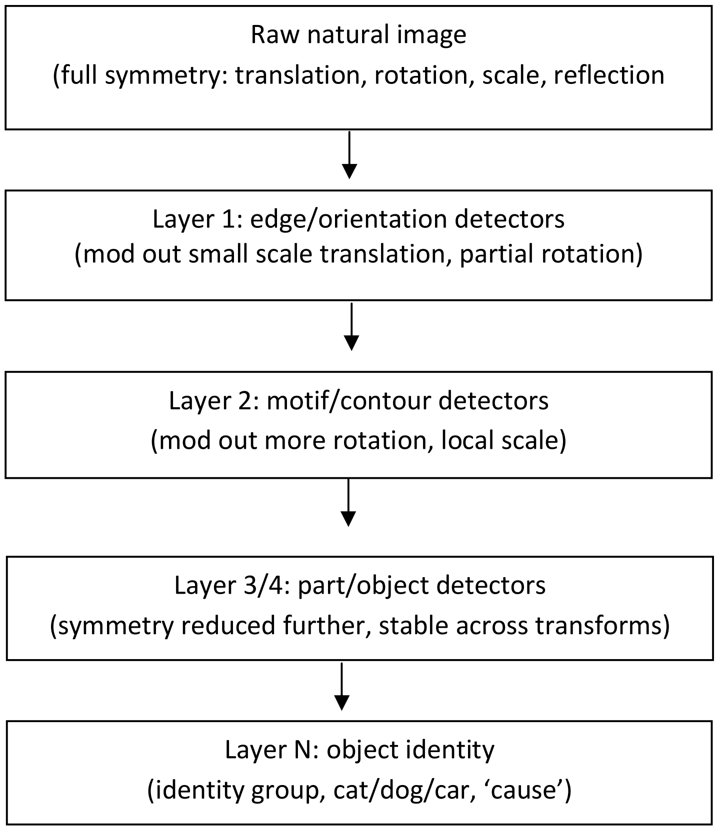


Figure 1 How a NN or the visual system processes data to identify the ‘causes’ of the data by removing symmetries at each layer.

A multilayer NN could take natural images, many of them, and reduce the redundancy at each layer. This can be thought of as a set of symmetries being removed at each layer. Layer after layer removes more and more symmetry that was present in the input ensemble, until there are no more symmetries to remove. The output layer has no symmetry left – all we have are causes (objects). These causes are the things that generated the images in the first place (figure 1).

We can think of this in term of group theory. The set of all symmetries contained in the ensemble of visual images is called the Root group. It can be very large. A sub-group is removed at each layer until at the final output there are no more symmetries left and all we have is the identity group. For instance oriented edges appear everywhere in a set of natural images. By identifying edges in little image patches and their orientations, with pooling of edges the network can say ‘edge thing somewhere in the image’. Edges can be combined into corners etc and further to objects and with the same pooling operations the NN output can say ‘cat somewhere in the image’.

**Symmetry in image data**

Images contain structured redundancies. If you photograph a tree, you can slide the picture a few pixels left (translation), rotate the camera slightly (rotation), or even mirror the scene (reflection), and you still recognize "tree."

For object recognition we would like to pick out things in images that remain the same regardless of whether it is scaled, rotated or translated. To do this we need to remove those symmetries. If we, or a neural network, does not do this then all it can do is store all the images it sees as sets of pixels in a giant lookup table and give them labels. New images would be compared to this lookup table and given a label that it most closely matches.  Nothing would have been learned from the inputs.

A machine learning system that doesn't somehow account for those symmetries wastes effort. Without symmetry handling, the network must relearn "treeness" separately for every shifted, rotated, or reflected version of the tree. Symmetry removal or "invariance building" means collapsing those infinitely many equivalent images into a smaller, manageable set of representations ('causes' which are objects in natural images).

**Local oriented filters as a starting point**

One symmetry to address is translation. The basic idea: instead of tying every neuron to every pixel in a fixed position (a fully-connected layer), you use local receptive fields. Each neuron looks at a small patch of pixels, applies a linear filter, and then pushes the result through a nonlinearity. Each little patch has a neuron that detects vertical lines and if all these neurons are connected to another neuron (in a higher layer) such that it is active when any of the inputs are active, the network has removed translation symmetry – there is a vertical edge somewhere in the visual field, but we don't know where. This is called pooling. We have just stripped out translation symmetry: shifting the image only shifts where the filter responds, not whether it responds.

**Why orientation matters**

Edges are important regularities of natural images. By learning a bank of oriented filters (e.g., horizontal, vertical, diagonal), the network builds a local code that describes any line-like feature in the patch. This is exactly what neuroscientists observed in the primary visual cortex: simple cells tuned to specific orientations. Now, why does orientation detection touch symmetry? Discarding orientation—by pooling across edges—means you get an approximately rotation-invariant representation (rotating the input leaves the response unchanged).

**Translation symmetry reduction step by step**

Take translation first. Imagine a vertical edge in the middle of an image. If you shift the image left by 10 pixels, the vertical edge is still there, just elsewhere. What happens across layers? We apply a max or average pooling operator across patches. Pooling discards the precise coordinates of where the edge fired, but preserves that it fired somewhere. After pooling, whether the edge was here or anywhere else doesn't matter. You've collapsed translations to the same representation.

Thus, the act of detecting oriented lines locally plus pooling across positions implements translation

**Rotation symmetry reduction step by step**

Rotation symmetry is trickier, but the principle is similar. Consider a vertical edge rotated by 30°. If you have only a vertical filter, a rotated edge won't activate it strongly. To cover rotations, you need a bank of filters at different orientations (say every 10°). A rotated edge will then activate whichever filter aligns with it. If you take the max (a neuron that will fire if any of the inputs is positive) across these orientation responses you throw away the precise angle information. Whether the edge was vertical, diagonal, or horizontal, the pooled value simply says "edge strength here." Once orientation invariance is built locally, higher layers can combine pooled edges into corners, curves, or more complex motifs. A corner formed from one vertical and one horizontal line remains a corner even if the whole figure is rotated 90°. Because lower layers already neutralized the orientation symmetry, the higher layer's detector becomes rotation-invariant. Thus, detecting oriented lines is the first step - pooling across them removes rotation symmetry.

The key point is that higher layers now just respond to objects such as trees. Trees, and other objects, that remain essentially those objects whether translated, rotated or scaled, are the causes of the data: natural images are made up of those objects. So by removing symmetries inherent in the original data, causes of the data can be identified.

**Symmetry reduction and redundancy reduction**

'Modding out' a symmetry group means collapsing all inputs related by that group into a single equivalence class. For instance, all the directed edges in an image patch can be collapsed into 'edge activity' by pooling. All the edges are in the same rotational symmetry group (SO(2)). The pooled output node is the representative of that symmetry class. Finding non-Gaussian directions (as in ICA) means identifying features that are statistically independent. Redundancy reduction or 'sparse coding' (Barlow (1961, 2001), Willmore et al (2011) Forte et al (2005), Simoncelli & Olshausen (2001), Chalk et al (2018), Vinje & Gallant (2000), Angelucci & Bressloff (2017)) is the computational goal of both — and it's what the brain seems to do in early sensory layers (Betramini et al (2018)).

Using NNs to find symmetry groups in the data has been discussed by Gabel (Gabel et al (2023)). Further discussion regarding symmetry and NNs can be found in (Armenta & Jodoin (2020), Gens & Domingos (2014), Laird et al (2025)).

**Biological evidence**

This layered strategy is not artificial. Neuroscience has shown similar behavior in the visual cortex: Simple cells in V1 respond to oriented lines at specific locations. Complex cells pool across positions, becoming translation invariant. Higher areas like IT cortex respond to objects regardless of rotation, scale, or position. Some NNs essentially use the same architecture, suggesting it may the natural solution to the symmetry problem in visual data. This process can be thought of in terms a subgroup chain with the Root group at one end (all symmetries contained in visual images) and the identity group at the other (figure 2).

**Subgroup chain of NN symmetries**

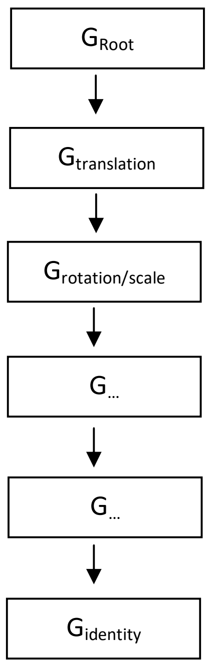


Figure 2 A general method of removing symmetry groups ($G_{translation}$, $G_{rotation/scale}$) in a dataset at each layer in a NN resulting in 'causes' ($G_{identity}$)

The brain is doing something like this but using all the data at its disposal, not just visual images but touch, hearing and so on. There will be a Root group (all the symmetries contained in the data a brain is exposed to) and a final group (the group of symmetries left over after everything that can be removed, is removed). We end up with 'causes' of the original data. We understand the world by uncovering these causes, which can be recombined to form versions of the datasets they were extracted from (the generative model).

What humans, through science, are trying to do is discover the laws of how the universe operates. As mentioned above, an important area that has been investigated is how symmetry plays a role in giving particles attributes, and the current best model (the Standard Model) is described by SU(3)xSU(2)xU(1) which refers to a combination of symmetry groups. The model can explain many observations about the universe such as why there are three ‘colours’ of quarks. Bigger symmetry groups have been proposed to account for further aspects of the physical world such as gravity.

The idea being put forward in this paper is that, other than doing science to uncover what the GUT symmetry group is, another approach would be to use a multilayer NN to process data, as much data as possible from as many sources as possible (particle accelerators etc.). The NN would automatically remove symmetry groups from the Root group. Then we could inspect the connections of the NN at each layer and, with the knowledge of the nonlinearities of each node, reconstruct the symmetry group removed at each layer. We only have to look at the connections, which we would have access to, to discover more about the laws of physics. We could construct a partial subgroup lattice (Schmidt (1994)) with the final goal of identifying the Root group, which would be the sought after GUT symmetry group.

A subgroup lattice in group theory is **a** partial order diagram showing how subgroups are nested within a group. Each edge **r**epresents **a** direct inclusion**:** if there's an edge from subgroup A to subgroup B, then B is a proper subgroup of A, and there's no subgroup strictly between them. A path through the lattice is a sequence of descending inclusions which in the case of $D_2$ (figure 3) would be:

$$D_2 \supset \langle s \rangle \supset \{e\}$$

But not

$$D_2 \supset \langle s \rangle \supset \langle r \rangle \supset \{e\}$$

Each step corresponds to modding out a subgroup — removing the symmetries it represents. The subgroup path that a NN could obtain would not necessarily be a path as defined above since it may mod out symmetry groups that are not connected by edges, and so instead we can call it a trajectory. This trajectory would contain subgroups that the Root group must contain. A single NN may not be able to identify the Root group because several candidate Root groups could contain those subgroups. However, if several different trajectories could be generated from the data, e.g. by a collection of NNs fed the data in different ways, then this may hone in on a candidate Root group. Mathematicians have knowledge about the subgroup lattices of many groups, especially candidate GUT groups like SU(5), SO(10) or E8, and so partial lattices found by NNs can be matched against candidate subgroup lattices. The idea would be to ‘cover’ the lattice sufficiently to identify the Root group.

In theory, if you start with a dataset whose structure is governed by a rich Root symmetry group G, and you train a well-constructed multilayer neural network that uses, for instance, nonlinear ICA at each layer (Hyvarinen et al (2018), Shan et al (2006)), then a careful spectral analysis of the weight matrices (eigenvalues and eigenvectors) with, as mentioned above, knowledge of the nonlinearities of each node, can reveal the **s**ymmetry subgroups removed at each layer**.** Different symmetry groups leave distinct spectral fingerprints in the eigenstructure of neural network weight matrices. By analyzing eigenvalues and eigenvectors, you can often infer which symmetries have been preserved or modded out. This gives you a powerful tool for reverse-engineering the network’s internal representation geometry.

**A toy Example**

Consider the dihedral group $D_2$, which encodes the symmetries of a rectangle — including 2 rotations and 2 reflections. When analyzing particle configurations invariant under $D_2$, we can trace a trajectory of symmetry reduction through its subgroup lattice. This process, known as "modding out" a symmetry group, collapses configurations related by group actions into equivalence classes, removing extrinsic redundancy. Starting with the full group $D_2$, we first mod out the cyclic subgroup ⟨r⟩ (rotation by 180 degrees), treating all rotated configurations as equivalent. This yields a 'quotient space' (a group theory term for describing the symmetries left after removing one of subgroups) where only reflections (vertical ⟨s⟩ and horizontal ⟨sr⟩) remain. Next, we mod out a reflection subgroup, such as ⟨s⟩, further collapsing configurations that differ by mirror symmetry. Continuing this process by removing ⟨sr⟩ we eventually reach the trivial group {e}, where all symmetries have been removed. The resulting representation reflects only the intrinsic structure of the data, potentially exposing its generative elements.

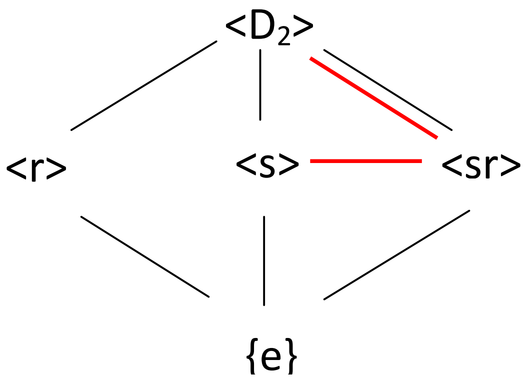


Figure 3 The subgroup lattice of $D_2$ (symmetries of a rectangle), with ⟨r⟩ a rotation by 180 degrees, ⟨s⟩ a horizontal reflection and ⟨sr⟩ a vertical reflection. A particular NN may follow the path in red indicating that the Root group must contain the two reflection symmetries. Another NN may only mod out ⟨r⟩ showing that, combined with the information from the first NN, that the Root group must contain the subgroups ⟨r⟩, ⟨s⟩, and ⟨sr⟩ suggesting $D_2$ as the possible Root Group.

In physics, this mirrors the journey from composite particle behaviour to fundamental constituents like quarks. In representation learning, it parallels the goal of disentangled, symmetry-reduced latent spaces. The subgroup lattice of $D_2$ provides a visual roadmap for this reduction, with each node representing a subgroup. Modding out all symmetry subgroups to reach the trivial group {e} removes all transformational redundancy, leaving the most invariant representation. Quarks, in this analogy, aren't literally {e}, but represent the irreducible generative elements — the fundamental structure that remains once all extrinsic symmetries have been stripped away.

However, since the Root group may be a Lie group, then there would be an infinite number of subgroups and it would be impossible to find all of those subgroups with a finite number of finite layered NNs. What can be done is to find parts of the lattice, and, by combining together the results from different NNs fed with different data, hope to cover enough of the lattice to identify the Root group.

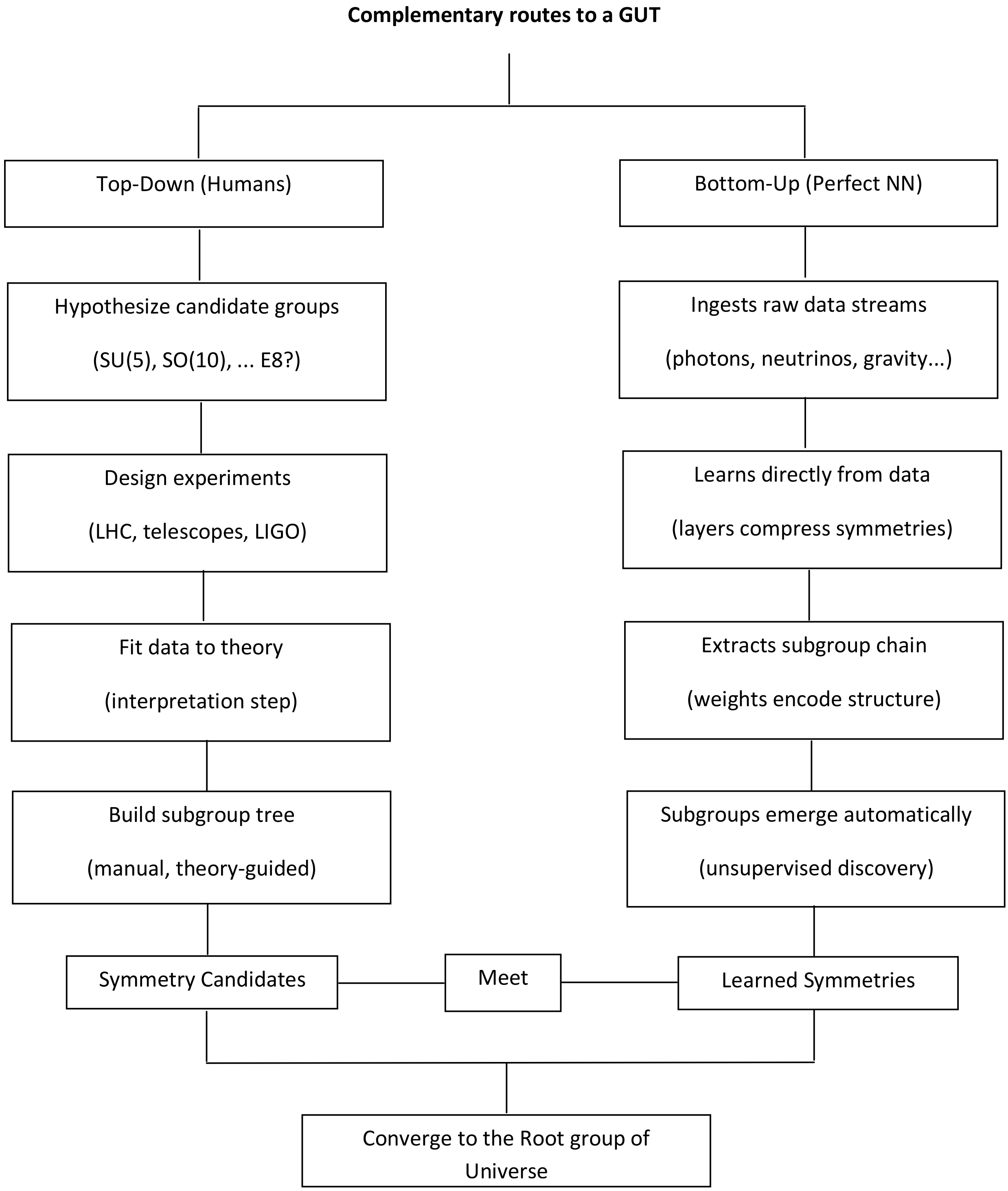


Figure 4 Complementary pathways using human brains and large NNs to discover the Root group of the Universe

**A plan for assisting scientists in finding the GUT symmetry group**

Suppose G is the ultimate Root Group. Train several large multilayer NNs on raw data streams. For each NN examine the eigenvalues and eigenvectors of the weight matrices at each layer. Together with knowledge of the node nonlinearities, deduce the symmetry group removed at each layer and construct a partial trajectory through the subgroup lattice of G for each NN. Match up the trajectories with known subgroup lattice structures for all candidate groups. Hopefully only one of the candidate group's subgroup lattice will align with the discovered partial lattice, and that group will be G (figure 4).

**References**


Angelucci, A., & Bressloff, P. C. (2017). *Circuits and mechanisms for surround modulation in visual cortex.* Annual Review of Neuroscience, 40**,** 425–451. https://doi.org/10.1146/annurev-neuro-072116-031418

Armenta M. A. & Jodoin,P (2020) *The Representation Theory of Neural Networks*, https://doi.org/10.48550/arXiv.2007.12213

Barlow, H. (1961) *Possible Principles Underlying the Transformations of Sensory Messages*. Sensory Communication 1 pp 217-234 10.7551/mitpress/9780262518420.003.0013

Barlow, H. B. (2001). *Redundancy reduction revisited*. Network: Computation in Neural Systems, 12(3), 241–253. https://doi.org/10.1080/net.12.3.241.253

Bertamini, M., Silvanto J., Norcia A.M., Makin A.D.J., and Wagemans, J. (2018) *The neural basis of visual symmetry and its role in mid- and high-level visual processing*. Ann N Y Acad Sci. 2018 Mar 31. doi: 10.1111/nyas.13667

Chalk, M., Marre, O., & Tkačik, G. (2018). *Toward a unified theory of efficient, predictive, and sparse coding*. Proceedings of the National Academy of Sciences, 115(1), 186–191. https://doi.org/10.1073/pnas.1711114115

Croon D., Gonzalo T. E. , Graf L. , Košnik N. , and White G. (2019) *GUT Physics in the Era of the LHC.* Frontiers in Physics Volume 7 - 2019 DOI=10.3389/fphy.2019.00076

Forte J.D., Hashemi-Nezhad M., Dobbie W.J., Dreher B., and Martin P.R. (2005) *Spatial coding and response redundancy in parallel visual pathways of the marmoset Callithrix jacchus*. Vis Neurosci. 2005 22(4):479-91. doi: 10.1017/S0952523805224094. PMID: 16212705.

Gabel*, A., Klein*, V., Valperga*, R., J.S.W. Webster, K. Quax, R. and Gavves, E. (2023) *Learning Lie Group Symmetry Transformations with Neural Networks* https://doi.org/10.48550/arXiv.2307.01583

Gens, R., & Domingos, P. (2014). *Deep Symmetry Networks*. Advances in Neural Information Processing Systems (NeurIPS) 3:2537-2545

Hyvärinen, A. & Hoyer, P.O. (2001) — A Two-Layer Sparse Coding Model Learns Simple and Complex Cell Receptive Fields. Vision Research Volume 41, Issue 18, August 2001, Pages 2413-2423

Hyvarinen , A. Sasaki , H. and Turner, R.E. (2018) *Nonlinear ICA Using Auxiliary Variables and Generalized Contrastive Learning*, https://doi.org/10.48550/arXiv.1805.08651

Kibble, T. (2014) *The Standard Model of Particle Physics*. European Review. 23. 10.1017/S1062798714000520.

Laird, L., Hsu, C., Bapat, A. and Walters, R. (2025) *MatrixNet: Learning over symmetry groups using learned group representations*. https://doi.org/10.48550/arXiv.2501.09571

Schmidt, R. *Subgroup Lattices of Groups*. (1994) de Gruyter Expositions of Mathematics, Vol. 14. Berlin–New York: de Gruyter

Shan, H., Zhang, L. and Cottrell, W.T. (2006) *Recursive ICA*. Advances in Neural Information Processing Systems: 19:1273-1280

Simoncelli, E. P., & Olshausen, B. A. (2001). *Natural image statistics and neural representation*. Annual Review of Neuroscience, 24**,** 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Vinje, W. E., & Gallant, J. L. (2000). *Sparse coding and decorrelation in primary visual cortex during natural vision*. Science, 287(5456), 1273–1276. https://doi.org/10.1126/science.287.5456.1273

Willmore, B.D.B., Mazer, J.A., and Gallant, J.A. (2011) *Sparse coding in striate and extrastriate visual cortex* Journal of Neurophysiology 2011 105:6, 2907-2919

Yang, J., Dehmamy, N., Walters, R. and Yu, R. (2024) *Latent Space Symmetry Discovery* https://doi.org/10.48550/arXiv.2310.00105