doi:10.20944/preprints202208.0402.v1

Fitness, Optima, and Simplicity

Kamaludin Dingle[1, 2, 3, *]
[1] Department of Computing and Mathematical Sciences, California Institute of Technology, USA
[2] Department of Chemical Engineering and Biotechnology, Cambridge University, UK
[3] CAMB, Department of Mathematics and Natural Sciences,
Gulf University for Science and Technology, Kuwait

Evolution by natural selection is often viewed as an optimisation process where an organism's phenotypic traits are adapted gradually to improve its fitness. Because of the many different traits with potentially conflicting requirements, among other factors, this optimisation process may appear onerous. Building on recent mathematical work connecting optima and simplicity, we here show that for certain generic phenotype fitness requirements — those based on physics and engineering principles — optimal phenotypic shapes will tend to 'simple', in the sense of low algorithmic or descriptional complexity. As a result, we argue that adapting to these types of generic fitness requirements will be a much 'easier' task for natural selection, compared to a null expectation based on arbitrary optimisation requirements. Further, selection's task may be easier still due to the fact that optimal phenotypes for one set of generic fitness constraints may also be close to optimal for other generic constraints, such that adapting to one constraint yields the other 'for free'.

Keywords: evolution; fitness; optimisation; algorithmic information theory

I. INTRODUCTION

Evolutionary adaptation is often viewed as an optimisation problem [1, 2], where the various traits of an organism are adjusted or 'tinkered' via natural selection, gradually improving fitness levels. Because organisms consist of very many traits, at many levels, evolutionary optimisation via searches in high-dimensional genome space may be regarded as a difficult task [3]. This is especially true if there are trade-offs between optimising different traits, or historical contingencies and developmental biases which limit accessible phenotypes [4].

Many aspects of an organism's fitness are environment-specific, for example an organism may evolve the ability to digest some newly encountered food source, but this ability is of little fitness value in the absence of that particular source. As another example, traits that increase the fitness of bacteria at high antibiotic concentration levels often decrease the fitness in the absence of the antibiotic [5], illustrating environment-specific trait adaptations.

However, there are also many aspects of fitness that are only weakly dependent on the specific environment in which an organism inhabits, in particular when trait fitness derives mainly from generic physics and engineering principles. For example, for a molecule to have a high level of robustness to genetic mutations, which is a generic fitness requirement for most organisms [6], is a property dependent on chemistry and physics laws. Also, mechanical stability is a generic fitness requirement for plants (among other organisms), and the constraints imposed by this requirement derive mostly from physics and engineering considerations. These kinds of generic fitness requirements are only weakly dependent on the organisms specific environment.

In this letter we will focus on adaptation in terms of optimisation with respect to these generic fitness requirements. Further, we will confine ourselves to studying phenotype morphologies, or shapes, and their levels of optimality with respect to these requirements. That is, we will assume that there is a large number of possible phenotypic shapes each affording different fitness benefits to an organism. From these shapes, we assume that natural selection will act to find highly optimal shapes that afford a high level of fitness benefit, and that the benefit derives from generic fitness requirements. In this setting, building on recent algorithmic information theory work in optimisation [7] and evolution [8], we will here argue for a 'fit is simple' claim, and also that these types of simple and 'fit' phenotypes are easy to adapt to. Hence evolutionary optimisation may be easier than it might at first appear, for the types of problems we study here.

II. THEORY

Optima and simplicity: We will begin by briefly reviewing some relevant theory and then follow this by demonstrating how the ideas relate to fitness and optimisation in evolution.

Algorithmic information theory (AIT) [9–12] concerns the information content of mathematical objects, such as numbers, patterns, graphs, discrete geometries, shapes, sequences, etc. In AIT, information content is defined by *Kolmogorov complexity*, which quantifies the amount of information required by a program, algorithm, or set of instructions to describe an object. Hence Kolmogorov complexity is also known as *descriptional complexity.* Simple patterns like $x = ababababababababababababababab$ can be described via short programs (eg "print ab 15 times"), while typical irregular patterns do not admit such brief and succinct descriptions, and are hence deemed to have high descriptional complexity. Measuring information content with Shannon entropy is well known in evolutionary biology [13], and although descriptional complexity is related to Shannon entropy, it differs in that the latter is only defined with respect to some probability distribution.

---

[*] dinglek@caltech.edu

A mathematical result [7] connecting optima and simplicity, which this work builds on, is outlined now:

(1) Assume we have a large collection $\mathcal{X}$ of different possible shapes (ie discrete geometries, configurations, patterns, sequences, etc) and some objective function $f(x)$ which is to be maximised over the elements $x \in \mathcal{X}$. If the function $f(x)$ is 'simple' in the sense of low descriptional complexity, then the optimal shape x^* must be 'simple' (as well as other close to optimal shapes).

(2) Assuming the same conditions as in (1), if some shape x^* is optimal for $f(x)$, then it is more likely to also be highly optimal for a different simple function $g(x)$, as compared to what would be expected for typical arbitrary functions.

See ref. [7] for details, a formal statement, and a proof. For the function $f(x)$ to have low descriptional complexity, we must be able to describe it precisely, succinctly, and unambiguously. Common laws of physics fall into this category [7]. In reference to (2), "expected for typical arbitrary functions" refers to the fact that if $f(x)$ and $g(x)$ were arbitrary functions, then for one shape to be highly optimal for both functions would be extremely unlikely, due to the large number of possible shapes in $\mathcal{X}$. The situation is very different when the functions $f(x)$ and $g(x)$ are not arbitrary, but instead are known to be 'simple' functions, in which case the probability may be much higher. The intuition behind these results is that "the optimal shape for function $f(x)$" is a very short description of the optimum x^*, and hence it must have low descriptional complexity, as long as the function $f(x)$ is also simple.

Simple optima in biological evolution: To connect now to evolution and adaptation, let us represent some set of possible phenotypic shapes with the collection $\mathcal{X}$. Assume that the fitness of an organism would be increased if the value of $f(x)$ increases, and hence that natural selection would search the elements in $\mathcal{X}$ for gradually more optimal phenotypes, with higher values of $f(x)$. In general, $f(x)$ may be a very complicated function taking into account all manner of environmental factors. In this case, it is hard to make universal statements about the properties and types of x that would result from selection. However, we propose that in biology, many — but certainly not all — fitness requirements derive from generic physics and engineering requirements. Examples of these generic fitness requirements include the need for thermal and mutational robustness in biomolecules, mechanical stability in plants, large surface areas for plants to capture sunlight, efficient information flow in biological networks, etc. These generic fitness requirements can often be described by a few equations, and are thus 'simple' in the AIT sense.

If we restrict ourselves to studying those aspects of fitness that rely only on generic physics and engineering constraints, then $f(x)$ will typically be a low complexity function. In this scenario then, we can invoke the optima-simplicity result outlined above in (1), and therefore infer that *highly optimal phenotypic shapes will have low descriptional complexity*, ie be 'simple'. Further, applying (2) now in our biological context, we infer a null model expectation that some phenotypic shapes that are highly optimal for one set of 'simple' constraints will be highly optimal for a different set of 'simple' constraints more often than would be expected for arbitrary constraint functions.

As a side comment, (1) can be seen to be related to the fact that because natural selection can be viewed as taking information from the environment and coding it into genome [14], it is intuitive that if the environmental constraints represented by $f(x)$ have low complexity, then the optimal phenotype x^* for the simple environment will also have low descriptional complexity. Result (1) can perhaps also be seen as a generalisation of the proposal that symmetry in organisms is a response to symmetric mechanical forces [15].

Note that low descriptional complexity does not always correspond to simple, regular, or symmetric forms [12]. Nonetheless, it was argued [7] that for the types of problems relevant in natural sciences, this correspondence can be expected to be commonly observed.

Optimal, simple, phenotypes may be easy to 'find': We have argued that fitter phenotype shapes for generic engineering-based requirements are expected to have low information content, and be 'simple'. Recently in ref. [8] we showed mathematically based on AIT arguments (Cf. ref. [16]) that during evolutionary searches in genome space, 'finding' short and simple programs is much easier than finding longer ones. Hence there is a *simplicity bias* [16] which biases evolution towards simpler, more symmetric, or regular morphologies (see also ref. [17]), and simple phenotypes are generally easier to adapt to. Earlier, Wolfram speculated that some of the simple and regular shapes in biology may be due to the fact that these require only short and easy to generate biological algorithms [18]. In brief, then, we claim here for a restricted class of adaptation problems, that optimal shapes are likely to be simple, and simple shapes are likely to be easy to adapt to.

Note that our claim 'fit is simple' is not the same as 'simple is fit'. There may be many simple phenotype shapes which are very suboptimal [7]. Hence the simplicity bias does not always bias the search towards more optimal shapes, but we argue that it will do so more often than expected for arbitrary constraint functions. This matter is related to that of low complexity, low probability shapes [19, 20].

Complex optima in biological evolution: There are very many cases in biology where simple phenotypes will not be more advantageous. If the environment is complex, and fitness does not depend primarily on generic physics based criteria, then phenotypic complexity will often tend to increase via selection [21]. As an example, Carothers et al [22] suggested that there might be a general positive relationship between the complexity of RNA aptamer structures and their functional activity, ie fitness. However, because an aptamer's functional activity is largely determined by how well its shape binds to some other specific molecular shape, this shape requirement is evidently not a generic requirement (to which our arguments apply), but instead a highly environment-specific one. Hence this apparent counterexample to our claim that 'fit is simple' in fact only illustrates a case where the environment is complex, leading to a complex optimal shape.

III. SUPPORTING EXAMPLES IN THE LITERATURE

We now survey several examples from the literature that lend support to our hypothesis that 'fit is simple'. Additionally, some of these examples also illustrate cases where shapes simultaneously are optimal, or close to optimal, for more than one fitness-related constraint function.

Lattice proteins: Lattice proteins are a well-studied and idealised model of protein folding. A protein is represented as a chain of balls and links, which is restricted to sit on a rigid lattice. It is known that the most designable proteins (ie those absorbing the largest number of possible genotype sequences) have simple and regular shapes [23, 24], and at the same time are fast folders [25], fast unfolders [26], thermally stable [23], highly evolvable [6], and highly mutationally robust [27]. Relatedly, in a simple continuum model of protein folding, kinetic optimisation was linked to symmetry of ground state fold [28]. So we see that simplicity is related to several generic physics-based fitness-related optimal properties.

Natural proteins: In natural protein tertiary structures, designability and thermal stability are also linked [29]. Due to the connection with thermal stability, England et al [30] point to designability as an important component of protein fitness. Further, significant associations among structural designability, folding rate and intrinsic disorder also exist [31]. Additionally, in ref. [32] it was reported that mutationally robust proteins tend to be more symmetric and also more compact.

Branching morphologies: Branching and fractal patterns are common in many organisms. Examples include fly trachea, mammalian salivary glands, and lung and kidney patterns [33]. These types of patterns are regular, repeating, and hence 'simple' in terms of descriptional complexity [34]. At the same time, such branching and fractal patterns are highly optimal for minimising energy loss and maximising space filling and surface area, which are biologically important [35] and hence related to fitness. Again we see examples of optimal shapes, which are also simple.

Cell lineages: Azevedo et al [36] suggest that simpler and regular metazoan cell lineages might develop faster, affording a fitness advantage. This suggests a beneficial fitness correlate of simplicity.

Plant morphology: Niklas [37] points out that the relative fitness of different plant morphologies (ie shapes) can be evaluated using basic physics or engineering principles. Important fitness constraints such as light interception, mechanical stability, ability to disperse spores, and conserving water are all related to general physics principles. Niklas noted that optimal plant morphologies in computer simulations are "comparatively simple in general appearance".

Strategy: In Axelrod's well-known Prisoner's Dilemma game theory tournament examining the evolution of cooperation, the winning (ie 'fittest') strategy was Tit-for-Tat, which was described as being extremely simple [38, 39]. This is a noteworthy example, because a strategy is not a phenotypic 'shape' exactly, but nonetheless can be rationalised via our information arguments. Later, Novak and Sigmund [40] developed another strategy, Pavlov, which outperformed Tit-for-Tat, but this new strategy was also noted to be extremely simple. The authors observed that very simple strategies may be expected to perform well in a wide variety of situations.

Scale free biological networks: Many important networks in biology have scale-free structures. Because of their ordered architectures, these networks have relatively low descriptional complexity. At the same time, they have many features that are positively linked to fitness-related properties of the networks, including responding quickly to external changes [41], being highly efficient [42–44], homeostasis against random failure [45], enhanced signal processing speed, computational power, and synchronizability [46], high error tolerance [47], good performance in uncertain environments [48], stability when new nodes are added [49], and evolvability [50].

The preceding showed varied examples where 'fit is simple'. We suggest that our mathematical theory given above rationalises these cases of optimal phenotypic shapes and forms, and their connection to simplicity.

IV. DISCUSSION

We have proposed a theory that for certain phenotypic traits, especially shapes or forms, optimal traits with respect to generic physics and engineering based constraints will tend to be 'simple' in the sense of having low descriptional complexity. Because low complexity shapes are expected to be easier for evolutionary searches to generate, here we argued that evolutionary adaptation to such constraints presents a relatively 'easy' task for natural selection. Supporting our mathematical arguments, we surveyed published examples according with our 'fit is simple' hypothesis. Some of these examples also appeared to illustrate the finding of ref. [7] suggesting that, some simple shapes can be highly optimal for multiple constraint functions. This multi-benefit property may also aid in evolutionary adaptation, by yielding other fitness related benefits 'for free'.

Although our work here posits a relationship between some specific types of phenotypes and fitness constraints, outlining a general relationship between simplicity, complexity, and fitness in evolution is not straightforward. In ref. [8] it was suggested that simplicity bias at the molecular level may facilitate the evolution of complexity at higher levels, which might aid fitness increases. On the other hand, simplicity bias may presumably hinder the evolution of the aspects of fitness for which higher complexity molecular shapes are advantageous. See refs. [51–54] for more discussions on when morphological (a)symmetry affords selective advantages.

Algorithmic information theory arguments applied in this work have yielded somewhat surprising conclusions of interest to evolution theory. Future work should utilise this relatively unexplored information perspective on evolution and biology.

Acknowledgements: The author acknowledges financial support from the Gulf University for Science and Technology for two Seed Grants (grant numbers 263301 and 253571).

[1] J.M. Smith. Optimization theory in evolution. *Annual Review of Ecology and Systematics*, 9:31–56, 1978.

[2] R.M.N. Alexander. *Optima for animals.* Princeton University

Press, 1996.

[3] Artem Kaznatcheev. Computational complexity as an ultimate constraint on evolution. *Genetics*, 212(1):245–265, 2019.

[4] H.G. Spencer. Optimization in evolution, limitations of. In Neil J. Smelser and Paul B. Baltes, editors, *International Encyclopedia of the Social and Behavioral Sciences*, pages 10882–10887. Pergamon, Oxford, 2001.

[5] Suman G Das, Susana OL Direito, Bartlomiej Waclaw, Rosalind J Allen, and Joachim Krug. Predictable properties of fitness landscapes induced by adaptational tradeoffs. *Elife*, 9, 2020.

[6] A. Wagner. *Robustness and evolvability in living systems*. Princeton University Press Princeton, NJ:, 2005.

[7] Kamaludin Dingle. Optima and simplicity in nature. *Preprints*, 2022080330, 2022.

[8] Iain G Johnston, Kamaludin Dingle, Sam F Greenbury, Chico Q Camargo, Jonathan PK Doye, Sebastian E Ahnert, and Ard A Louis. Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution. *Proceedings of the National Academy of Sciences*, 119(11):e2113883119, 2022.

[9] R.J. Solomonoff. A preliminary report on a general theory of inductive inference. 1960.

[10] A.N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of information transmission*, 1(1):1–7, 1965.

[11] Gregory J Chaitin. A theory of program size formally identical to information theory. *Journal of the ACM (JACM)*, 22(3):329–340, 1975.

[12] M. Li and P.M.B. Vitanyi. *An introduction to Kolmogorov complexity and its applications*. Springer-Verlag New York Inc, 2008.

[13] C. Adami. The use of information theory in evolutionary biology. *Annals of the New York Academy of Sciences*, 2011.

[14] M. Kimura. Natural selection as the process of accumulating genetic information in adaptive evolution. *Genetical Research*, 2(01):127–140, 1961.

[15] Gábor Holló. Demystification of animal symmetry: Symmetry is a response to mechanical forces. *Biology Direct*, 12(1):1–18, 2017.

[16] Kamaludin Dingle, Chico Q Camargo, and Ard A Louis. Input–output maps are strongly biased towards simple outputs. *Nature communications*, 9(1):761, 2018.

[17] Pascal F Hagolani, Roland Zimm, Renske Vroomans, and Isaac Salazar-Ciudad. On the evolution and development of morphological complexity: A view from gene regulatory networks. *PLoS computational biology*, 17(2):e1008570, 2021.

[18] Stephen Wolfram. *A New Kind of Science*. Wolfram Media, 2002.

[19] Kamaludin Dingle, Guillermo Valle Pérez, and Ard A Louis. Generic predictions of output probability based on complexities of inputs and outputs. *Scientific reports*, 10(1):1–9, 2020.

[20] Mohamed Alaskandarani and Kamaludin Dingle. Low complexity, low probability patterns and consequences for algorithmic probability applications. *arXiv preprint arXiv:2207.12251*, 2022.

[21] Joshua E Auerbach and Josh C Bongard. Environmental influence on the evolution of morphological complexity in machines. *PLoS computational biology*, 10(1):e1003399, 2014.

[22] J.M. Carothers, S.C. Oestreich, J.H. Davis, and J.W. Szostak. Informational complexity and functional activity of RNA structures. *Journal of the American Chemical Society*, 126(16):5130–5137, 2004.

[23] H. Li, R. Helling, C. Tang, and N. Wingreen. Emergence of preferred structures in a simple model of protein folding. *Science*, 273(5275):666, 1996.

[24] T. Wang, J. Miller, N.S. Wingreen, C. Tang, and K.A. Dill. Symmetry and designability for lattice protein models. *The Journal of Chemical Physics*, 113:8329, 2000.

[25] R. Mélin, H. Li, N.S. Wingreen, and C. Tang. Designability, thermodynamic stability, and dynamics in protein folding: a lattice model study. *The Journal of chemical physics*, 110:1252, 1999.

[26] C.L. Dias and M. Grant. Designable structures are easy to unfold. *Physical Review E*, 74(4):042902, 2006.

[27] Sam F Greenbury, Steffen Schaper, Sebastian E Ahnert, and Ard A Louis. Genetic correlations greatly increase mutational robustness and can both reduce and enhance evolvability. *PLoS computational biology*, 12(3):e1004773, 2016.

[28] E.D. Nelson, L.F. Teneyck, and J.N. Onuchic. Symmetry and kinetic optimization of proteinlike heteropolymers. *Physical review letters*, 79(18):3534–3537, 1997.

[29] N.S. Wingreen, H. Li, and C. Tang. Designability and thermal stability of protein structures. *Polymer*, 45(2):699–705, 2004.

[30] J.L. England, B.E. Shakhnovich, and E.I. Shakhnovich. Natural selection of more designable folds: a mechanism for thermophilic adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15):8727, 2003.

[31] E. Ferrada and A. Wagner. Protein robustness promotes evolutionary innovations on large evolutionary time-scales. *Proceedings of the Royal Society B: Biological Sciences*, 275(1643):1595–1602, 2008.

[32] J. Hartling and J. Kim. Mutational robustness and geometrical form in protein structures. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 310(3):216–226, 2008.

[33] P. Lu and Z. Werb. Patterning mechanisms of branched organs. *Science*, 322:1506, 2008.

[34] R.J. Metzger, O.D. Klein, G.R. Martin, and M.A. Krasnow. The branching programme of mouse lung development. *Nature*, 453(7196):745–750, 2008.

[35] G.B. West, J.H. Brown, and B.J. Enquist. A general model for the origin of allometric scaling laws in biology. *Science*, 276(5309):122–126, 1997.

[36] R.B.R. Azevedo, R. Lohaus, V. Braun, M. Gumbel, M. Umamaheshwar, P.M. Agapow, W. Houthoofd, U. Platzer, G. Borgonie, H.P. Meinzer, et al. The simplicity of metazoan cell lineages. *Nature*, 433(7022):152–156, 2005.

[37] Karl J Niklas. Computer models of early land plant evolution. *Annu. Rev. Earth Planet. Sci.*, 32:47–66, 2004.

[38] Amnon Rapoport. Optimal policies for the prisoner's dilemma. *Psychological Review*, 74(2):136, 1967.

[39] R. Axelrod and W.D. Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.

[40] M. Nowak, K. Sigmund, et al. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(6432):56–58, 1993.

[41] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, 2000.

[42] E. Melendez-Hevia, T.G. Waddell, and F. Montero. Optimization of metabolism: the evolution of metabolic pathways toward simplicity through the game of the pentose phosphate cycle. *Journal of theoretical biology*, 166(2):201–220, 1994.

[43] R. Cohen and S. Havlin. Scale-free networks are ultrasmall. *Physical Review Letters*, 90(5):58701, 2003.

[44] V. Colizza, J.R. Banavar, A. Maritan, and A. Rinaldo. Network structures from selection principles. *Physical review letters*, 92(19):198701, 2004.

[45] S. Valverde, R.F. Cancho, and R.V. Sole. Scale-free networks from optimal design. *EPL (Europhysics Letters)*, 60:512, 2002.

[46] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[47] R. Albert, H. Jeong, and A.L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

[48] J.M. Carlson and J. Doyle. Highly optimized tolerance: Robustness and design in complex systems. *Physical Review Letters*, 84(11):2529–2532, 2000.

[49] J.I. Perotti, O.V. Billoni, F.A. Tamarit, D.R. Chialvo, and S.A. Cannas. Emergent self-organized complex network topology out of stability constraints. *Physical review letters*, 103(10):108701, 2009.

[50] S.F. Greenbury, I.G. Johnston, M.A. Smith, J.P.K. Doye, and A.A. Louis. The effect of scale-free topology on the robustness and evolvability of genetic regulatory networks. *Journal of Theoretical Biology*, 267:48–61, 2010.

[51] Anders Pape Møller and John P Swaddle. *Asymmetry, developmental stability and evolution*. Oxford University Press, UK, 1997.

[52] Sebastian Ocklenburg and Annakarina Mundorf. Symmetry and asymmetry in biological structures. *Proceedings of the National Academy of Sciences*, 119(28):e2204881119, 2022.

[53] Iain G Johnston, Kamaludin Dingle, Sam F Greenbury, Chico Q Camargo, Jonathan PK Doye, Sebastian E Ahnert, and Ard A Louis. Reply to ocklenburg and mundorf: The interplay of developmental bias and natural selection. *Proceedings of the National Academy of Sciences*, 119(28):e2205299119, 2022.

[54] Edward Bormashenko. Fibonacci sequences, symmetry and ordering in biological patterns, their sources, information origin and the landauer principle. *Preprints*, 2022080107, 2022.