Interferometric “fitness” and the large optical array

David Buscher and Chris Haniff

Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

ABSTRACT

We introduce for the first time an approach to interferometer design which makes explicit the global nature of the optimisation of the array. We discuss the idea of a ”fitness landscape” for interferometric arrays, where a fitness function is defined in terms of the success of the array, and the coordinates of the space are design parameters such as numbers and sizes of telescopes, baseline length and wavelength coverage. We investigate the extent to which such ideas can illuminate our understanding of strategies for array design and our ways of evaluating the success of a design.

Keywords: Interferometer design, interferometric array optimisation

1. INTRODUCTION

A modern optical or infrared interferometer is a complex device consisting of a large number of components such as telescopes, delay lines and beam combiners, each of which are in themselves complex devices. The task of array design is to choose a set of components and parameters for these components in such a way that they all work together to produce the best possible interferometer for a given purpose. This paper discusses principles which need to be taken into account when evaluating array designs to decide which design is the “best”.

Our discussion makes extensive use of the idea of “fitness landscapes”, an idea developed originally in evolutionary biology. We view array design as analogous to the “design” of a biological organism, i.e. that of making a complex system which works. Although the method of solution of the these two design problems is different, in one case evolutionary and in the other (hopefully) deliberate, the nature of the problem is similar in the two cases. This cross-disciplinary approach exposes a number of truths which are common to all types of system design and serves to reinforce our understanding of these truths in the context of interferometric arrays.

We begin by reviewing the idea of a fitness landscape as developed in a biological context. We then show how this idea can be translated into an interferometric context and show how application of these concepts leads to a number of specific results.

2. FITNESS LANDSCAPES

The idea of fitness landscapes was first developed by Sewald Wright[1] as an aid to understanding the processes by which organisms in nature develop into different species. The idea is a very simple: we imagine a set of organisms with N genes as populating an N-dimensional space where each coordinate represents the value of one gene, i.e. the genotype is encoded as a position in this space. Associated with every genotype is a phenotype (i.e. physical expression of the set of genes as an organism) and for each phenotype we can define an evolutionary “fitness” which is related to the organism’s ability to survive and flourish. Wright quantified this in terms of the probability that the organism would survive and reproduce itself, and developed the idea of a landscape where the “height” of the surface was the fitness associated with a given coordinate in genotype space.

This landscape is likely to be a very complex and multi-dimensional one but nevertheless one can gain an intuitive feel for it by imagining a landscape defined on a two-dimensional subspace of the genotype space and relating features in this landscape to familiar features in geographical landscape, such as plains, mountains, valleys and so on. This metaphor serves an intuitive way of understanding the *structure* of complex systems, and has been applied in a number of fields from immunology to computer science.

Send correspondence to: David Buscher. E-mail: dfb@mrao.cam.ac.uk, Telephone: 44 1223 337302

3. DEFINITIONS IN AN INTERFEROMETRIC CONTEXT

To understand the utility of fitness landscapes to interferometry, we first need to make more explicit the analogy between interferometric arrays and biological organisms. The "genes" of an array we define as the set of properties of an array which specify its design completely, for example the number of telescopes in the array, their diameter, the array layout, whether or not a dual-feed scheme is employed, the beam-combining scheme and so on. Clearly this is a very high-dimensional space, although some dimensions of the space will have only a small number of options associated with them, and some dimensions will be much more important to the performance of the array than others. Nevertheless it is possible to imagine a finite-dimensional space which encompasses the majority of useful interferometers, as well as a near infinity of useless interferometers — we will discuss this aspect later.

More controversial is the choice for the "fitness" measure for an interferometer. One can think of a number of partial measures of an array's success, for example the number of measurements the array is capable of making or their accuracy. We choose here to stay relatively close to the spirit of a Darwinian approach: we define the "fitness" of an array (or more strictly speaking, an *array design*) in terms of the likelihood that the array will attract adequate funds for its construction and continued operation and development. The probability of attracting sufficient funding amounts to a "survival probability" for an interferometer and is thus directly equivalent to the macroevolutionary idea of fitness. This measure has the advantage of being a value-neutral one: it does not of itself pre-judge what the aim of an interferometer is (for example astrometry, imaging or high-precision visibility measurements), but like the notion of the "selfish gene" it defines the aim of an interferometer as simply to survive. A high-fitness array is more likely to be funded and therefore more likely to achieve any subsidiary goals such as doing science or technology development than a low-fitness array which struggles to get funding, and thus fitness is a good indicator of general success.

We can recognise a sequence of relationships between an interferometer design and its ultimate fitness. The array design (the genotype) is expressed as an actual array (the phenotype) which will have certain characteristics such as a cost and timescale to build, a limiting magnitude, resolution, wavelength coverage and so on. These characteristics in turn affect the performance of the array in terms of number of scientific papers produced, its public profile and its acceptance in the scientific community. This affects the perceptions of the funding agencies and other potential donors and hence the overall probability of funding i.e. the fitness. We can see these relationships as a sequence of "mappings" which map from the genotype space onto a unique fitness measure.

This definition of fitness is open to a number of criticisms, the most obvious being that this measure is difficult to quantify, since one can only really judge whether an array has been successful in attracting funding after it has been built. At that point it is too late to change the design, and what is really wanted in any case is the relative *probability* that a design will get funding, all other things being equal. Another difficulty in making this measure quantitative is the question of deciding what is as "adequate" funding for an interferometer to survive. Nevertheless, this definition of fitness serves as a good starting point for qualitative discussions of array performance because of its generality (we will not have to change our basic definitions when discussing different types of arrays) and its objectivity (different observers are unlikely to strongly disagree as to whether an array has survived or not).

In this approach we optimise something which we definitely want but find difficult to quantify. This is in contrast with approaches in which what is optimised is some aspect of an array, such as the signal-to-noise ratio for making an image of a point source,[2] which is easy to quantify but does not adequately represent what we want — we don't want to make images of point sources! One can hope with a general approach such as this one to unify our definition of "success" while in time becoming more sophisticated in our understanding of the strategies to achieve this success: most biologists still agree with Darwin's idea that the ability of an organism to survive and reproduce is the principal factor in natural selection, but current evolutionary thinking emphasises the fact that brute force and size do not necessarily maximise an organism's evolutionary fitness, and more subtle effects such as social behaviour can have a much larger effect on a species' ability to flourish.

The idea of an interferometric fitness landscape then forms a framework in which to discuss the design of arrays. We can look at the structure of the landscape, see how the "peaks" are arranged in this space, and from this draw conclusions as to what aspects of array design are most important and how to go about building the best possible interferometric array.

4. INSIGHTS FROM THE ANALOGY

Viewing array design from the standpoint of biological evolution allows us to see aspects of array design which might be less obvious from more conventional standpoints. A large range of different ideas result from this, but the ideas we will discuss in the rest of this paper fall under three main headings, summarised as follows:

Fitness is a global measure in a multidimensional space. The success of an array is unlikely to come from a single aspect of a design but rather from the combination of a number of features chosen in the right proportions to complement each other.

Fitness is always defined with respect to an environment. There are many factors external to the array design which affect the success of the array, which we call here the "environment". Array design can be seen as the process of adapting to this environment in an optimal way.

The fitness landscape is inherently "rugged". There are likely to be multiple local peaks in the landscape separated by areas of low fitness and therefore an incremental change from a local peak is likely to lead to a reduction in fitness.

We discuss these ideas in some detail in the following sections, giving examples of the applications of these ideas to particular aspects of interferometer design.

5. GLOBAL OPTIMISATION IN A MULTIDIMENSIONAL SPACE

It is not hard to recognise that stating the problem of array design in terms of fitness landscapes is equivalent to stating the problem as one of multidimensional optimisation. A number of general results from the field of optimisation are therefore immediately applicable. A trivial result from this field is that the optimum is unlikely to be reached by searching along only one axis in the multidimensional space. This is equivalent to saying that the optimum generally comes from having the right balance of parameters in a design.

An example taken from evolutionary biology is that a predator does not simply optimise its speed so that it can catch prey, it must also develop the teeth and strength so that it can bring the prey down — a toothless but very fast predator may lose out to a slower predator with teeth. In a similar way an array design which only optimises a single factor such as angular resolution is unlikely to be as successful as one that has a better balance of angular resolution with, for example, sensitivity. Bringing together the right mixture of all the parameters of an array so that they work together is how we get the best possible scientific results from the array.

The key idea is that we should not confuse optimising any of the intermediate variables in the mapping from genotype to fitness (such as signal-to-noise ratio or number of observations per night) with optimising the fitness itself.

6. FITNESS AND THE ENVIRONMENT

An important factor in the biological theory of fitness landscapes is that the landscape can depend on the environment in which an organism finds itself: the fitness value associated with a given coordinate in genotype space can be very different depending on the environment in which fitness is being measured. For example, the fitness landscape clearly has a higher peak for the genotype of a polar bear than for that of a camel when the environment is the arctic whereas the fitness of the camel genotype is clearly higher when the environment is a desert. The organism does not optimise in a vacuum but adapts itself to the environment it is in. Note that the "environment" in this case does not refer to a local region of genotype space but rather to something which is *outside* the genotype space altogether — the genotype defining a polar bear does not depend on whether it is in the arctic or in the desert.

The equivalent for the "environment" in the interferometric fitness landscape is anything which is outside the control of the design but which affects the fitness of the design. There are many such factors: for example, we can define a "source environment" as the parameters of the set of astronomical sources that are available to be observed. There is nothing an array designer can do to change the angular sizes and brightnesses of these sources, but if an interferometer can only observe a certain class of source and no such sources exist, the array is unlikely to be judged a success. For example, the nearest star-forming regions are all more than 100 parsecs away, and this implies that interferometers with an angular resolution poorer than 10 milliarcseconds will be unable to image any planet-formation processes on AU-type scales — we cannot bring the star-formation regions any closer to compensate for this!

We can also conceive of a "scientific environment" which determines what classes of observations are seen as interesting or necessary to advance the progress of science. For example, angular diameter measurements of Cepheids are currently felt to be more valuable than angular diameter measurements of solar-type stars.

The list of such factors could go on to include the funding environment (which determines the likelihood that an expensive array will get built) and many other factors, but the point is that it is an important part of array design to recognise the nature of the environment the array must operate in and to adapt the array to its environment in the same way that a biological organism adapts itself. One can always try to influence some aspects of the environment (in the same way that tree roots break through rocks to make a better environment for collecting nutrients) but it is a fatal mistake to pretend that the environment is not there.

7. RUGGED LANDSCAPES

An important question in evolutionary biology is that of differentiation of species. Natural selection causes organisms to move towards the points of highest fitness, so one can ask why organisms have not all evolved to the same point on the landscape, i.e. why organisms do not all share the same general characteristics. One of the explanations for this is the "ruggedness" of the fitness landscape, which is to say that the fitness landscape consists of many isolated peaks of high fitness separated by valleys of much lower fitness. In such a landscape two organisms starting from nearby points may evolve in such a way that they are going "uphill" in the fitness landscape all the time but eventually reach two distinct "peaks", i.e. different species. Once at these peaks, the organisms will be unlikely to "cross over" to peaks with even higher fitness because their evolutionary paths would have to traverse the valleys in between and thus go through a stage of low fitness. Thus species represent "islands" of high fitness separated by a "sea" of low fitness.

The reason that the biological fitness landscape is rugged is called "epistasis", which is due to the fact that different genes interact with each other: the fitness afforded by one gene can be radically altered by a change in another gene. A genetic mutation which may cause an increase some desirable property in the organism may therefore cause a fatal decrease of some other desirable property. Thus the range of changes to genes which result in a net increase in fitness is restricted, and local peaks appear in the landscape.

There is some evidence that the interferometric fitness landscape is equally rugged because interferometers with quite different characteristics have been successful at the same time. For example aperture masking experiments with short baselines but a large number of apertures have been successful,[3,4] as have interferometers such as the PTI with a small number of apertures but much longer baselines.[5] The cause of this ruggedness is less clear until one takes funding limitations into account. An interferometer with long (and short) baselines and large numbers of apertures might have been able to dominate scientifically over either of the interferometer classes mentioned above, but this would have probably been too expensive. Thus funding limitations cause many of the design parameters of an array to interact with each other, and many "species" are able to survive separately.

One consequences of having a rugged landscape is that we need to decide in advance which "peak" in interferometric space we want to be on, because we cannot simply evolve from one peak to a better one. It may be the case, for example, that an interferometer optimised for imaging science cannot be converted later to do astrometry because that would involve evolution through extremely "unfit" space: having an interferometer which is good at both imaging and astrometry could be either impossible or too expensive.

Another consequence of ruggedness is that the best "peak" may be different from any of the peaks that other interferometers are currently sitting on. Relatively small changes of a number of different parameters of the design may mean that we move to a completely different "island" of success, perhaps doing different types of science entirely.

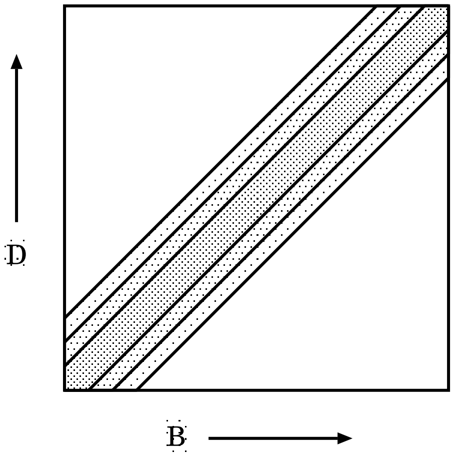

Figure 1. A schematic contour plot of the fitness landscape for an interferometer with two degrees of freedom: B the interferometer baseline and D the aperture diameter. The region of peak fitness (denoted by the darkest regions) is a "ridge" along a line of constant D/B.

8. PRACTICAL EXAMPLES

In this section we give some examples which illustrate the insights from fitness landscapes in more detail.

8.1. The "dilutedness ridge"

Here we take an example of a simplified design problem in which there are only two "genes" to our interferometer, so the fitness landscape is 2-dimensional. We imagine the design of a 2-element interferometer in which only the aperture diameter D and the maximum baseline length B are free parameters. Having a larger baseline will increase the angular resolution and having larger apertures will (in the absence of atmospheric seeing) allow us to see fainter sources. An increase of either will cost more money. What should we do?

Two principles we can apply here are multidimensionality and adaptation to the environment. It is no use considering aperture diameter independently of baseline length and it is foolish to ignore what sources are in the sky. If we are set the task, for example, of observing a class of sources of a fixed surface temperature T at a wavelength λ then the source environment will dictate a "ridge" in fitness space in which certain *combinations* of B and D are favoured.

To understand this we first note that the fixed temperature T means that the angular diameter of the source and its flux are related. The flux from of a source of angular diameter α is approximately $\alpha^2 P(T,\lambda)$ where $P(T,\lambda)$ is the blackbody Planck function. We can only do interesting interferometry on a source if it is approximately the same size as our angular resolution i.e. $\alpha \sim \lambda/B$ so the flux from the sources we typically observe will be $(\lambda/B)^2 P(T,\lambda)$. Thus the larger we make B the fainter the sources will be and *vice versa.* An interferometer with large B and small D will not be able to make any interesting measurements at all and so will have very low fitness, and conversely we will have wasted money on building an interferometer with large D but small B.

Thus there is a "ridge" in the (B,D) landscape where the interferometer fitness is maximised as illustrated in figure 1, a ridge of constant D/B i.e. of constant aperture "dilutedness", It should be noted that the existence of this ridge is due in large part to the nature of the astronomical sources and not to the technology of the interferometer, i.e. it is a source environment effect.

8.2. Imaging and the "scientific environment"

A matter of some controversy is the question of whether to do "precision interferometry" i.e. make very high accuracy measurements, typically of V^2, on a few baselines or to do "imaging", taken to mean collecting amplitude and (closure) phase measurements on enough different baselines to allow the reconstruction of model-independent images. Of course, one would ideally like to do both, but funding limitations may make this infeasible. We can give an answer to this question by relating it to fitness in the "scientific environment".

If we ask the question of what sources are the most interesting to observe with an interferometer, the answer is undoubtedly the sources for which information about the spatial structure of the emission on the angular scales only accessible by interferometers is critical to understanding the astrophysics. An example of this is the structure of dust tori in AGN, whose exact geometries and opacities are poorly constrained by spectrophotometric measurements.

This, then, is the scientific environment: interesting sources tend to be ones with structures on scales not accessible by other techniques. Consequently, the theoretical models of these sources on these angular scales is poorly understood because there is little other evidence to constrain them. Utilising precision V^2 measurements involves fitting these measurements to models, but these models will typically have many free parameters which are poorly understood. As a result, V^2 measurements are likely yield constraints only on a mixture of these parameters and say nothing about the overall correctness of the model. A model-independent image, on the other hand, can be directly compared with models to see whether it agrees: any differences between the images and those predicted by a model can be directly interpreted in terms of the correctness of the model.

One example of this is the case of IRC+10216, where spectral-energy-distribution fits are consistent with a spherical distribution of infrared emission[6] but where interferometric images[3, 7] clearly reveal a more complex geometry.

Thus V^2 and imaging measurements have different fitness because the scientific environment biases us towards observing sources which V^2 measurements are less suited to constraining. It is interesting to note that model-independent imaging has historically been one of the least-optimised features of interferometers. This is possibly because the increase in V^2 accuracy offered by some design feature is easy to quantify, and so has been optimised in isolation, to the detriment of the overall fitness.

8.3. Astrometry and planet detection

A topical example of how the environment of an interferometer can change rapidly is provided by the current focus on astrometric interferometry for planet detection. This has almost wholly been determined by the recent successes[8] of radial-velocity campaigns in detecting planets around solar-type stars. Prior to 1995 astrometric interferometry was seen as a rather narrow specialism, but since then it has become the focus of both the Keck and VLT interferometers.

9. FITNESS AND THE FUTURE OF INTERFEROMETRY

There are clearly a number of avenues open to exploration in the context of fitness landscapes in interferometry. One of the obvious questions is a deeper study of the question of ruggedness in the landscape: are the best interferometers special-purpose interferometers which are optimised for specific tasks, each living on its own "island", or is there a "mega-interferometer" which can do everything? Another question related to ruggedness is whether incremental changes to existing interferometers is an appropriate way to improve their fitness or whether they are already at a local peak in the fitness landscape such that the only route forward is to build new interferometers in entirely different regions of parameter space.

Another topic worth exploring is the idea of *viability*, by which we mean the question of whether there is some absolute level of fitness below which an interferometer is not worth building. Is there a "sea level" in the fitness landscape below which an interferometer will drown, and where are current interferometric designs in relation to that level?

Another question, common to the more general theory of fitness landscapes is that of dynamic landscapes: how does the landscape change with time and what effect does this change have on the fitness of existing interferometers? Clearly as new interferometers come on line, they will change the scientific environment by making new discoveries. This will mean that some classes objects will become less interesting because the interesting measurements have been made, so that older designs of interferometer will become obsolete. It will also (hopefully) mean that new areas of scientific interest will be opened up that were not considered interesting previously, so that new types of interferometers will be designed to explore these new regions of parameter space.

ACKNOWLEDGMENTS

The authors gratefully acknowledge inspiration from C. Brown who first brought their attention to the importance of landscapes, and encouragement from G. Ramsay and W. Kneissl.

REFERENCES

1. S. Wright, "The roles of mutation, inbreeding, crossbreeding and selection in evolution," in *Proceedings of the Sixth International Conference on Genetics*, **1**, pp. 356–366, 1932.
2. S. R. Kulkarni, S. Prasad, and T. Nakajima, "Noise in optical synthesis images. II. Sensitivity of an ${}^nC_2$ interferometer with bispectrum imaging," *Journal of the Optical Society of America* **8**, pp. 499–510, Mar. 1991.
3. P. G. Tuthill, J. D. Monnier, W. C. Danchi, and B. Lopez, "Smoke Signals from IRC +10216. I. Milliarcsecond Proper Motions of the Dust," *Astrophys. J.* **543**, pp. 284–290, Nov. 2000.
4. P. G. Tuthill, J. D. Monnier, and W. C. Danchi, "A dusty torus around the luminous young star LkHα101," *Nature* **409**, pp. 1012–1014, Feb. 2001.
5. G. T. van Belle, D. R. Ciardi, R. R. Thompson, R. L. Akeson, and E. A. Lada, "Altair's Oblateness and Rotation Velocity from Long-Baseline Interferometry," *Astrophys. J.* **559**, pp. 1155–1164, Oct. 2001.
6. Z. Ivezic and M. Elitzur, "Dusty Envelope Around IRC+10216," *Bulletin of the American Astronomical Society* **28**, pp. 943–+, May 1996.
7. C. A. Haniff and D. F. Buscher, "Variable sub-arcsecond structure in the circumstellar envelope of IRC+10216," *Astron. Astrophys.* **334**, pp. L5–L8, June 1998.
8. M. Mayor and D. Queloz, "A Jupiter-Mass Companion to a Solar-Type Star," *Nature* **378**, pp. 355–+, Nov. 1995.