Efficiency in neural information processing

Antti Knuutila
Master's Thesis
Cognitive Science
Institute of Behavioural Sciences
Faculty of Behavioural Sciences
University of Helsinki
2012

HELSINGIN YLIOPISTO - HELSINGFORS UNIVERSITET - UNIVERSITY OF HELSINKI

Tiedekunta - Fakultet - Faculty	Laitos - Institution – Department	
Faculty of Behavioural Sciences	Institute of Behavioural Sciences	
Tekijä - Författare - Author		
Antti Knuutila		
Työn nimi - Arbetets titel - Title		
Efficiency in neural information processing		
Oppiaine - Läroämne - Subject		
Cognitive Science		
Työn laji ja ohjaaja(t) - Arbetets art och handledare – Level and instructor	Aika - Datum - Month and year	Sivumäärä - Sidoantal - Number of pages
Master's thesis	August 2012	81

Tiivistelmä - Referat - Abstract

Brains are capable of processing information with remarkable efficiency under constraints set by the limited availability of physical resources such as the amount of space and the supply of metabolic energy. Natural selection has optimised the structure and function of brain networks using simple design rules similar to those found in man-made electronic and information systems. This study presents findings concerning a number of general principles of brain design governing the evolution and organisation of neural information processing.

The rule of minimising wiring in neuronal networks is one such principle operating on multiple levels of brain organisation. Both individual components and larger brain architectural units are seen to feature characteristics of near-optimal wiring. Miniaturisation of neuronal components conserves space but raises problems about noise in signalling. Small-world organisation of anatomical and functional networks is widely employed in the brain, contributing to high global efficiency at low cost. Metabolic costs severely constrain signal traffic in the human brain, necessitating the use of energy-efficient sparse neural representations.

Extensive evidence is presented of anatomical and physiological optimisations facilitating efficient information processing in brain networks. Limitations of current experimental techniques are discussed, with a view on possible future avenues of research.

Avainsanat – Nyckelord - Keywords

optimisation, efficiency, adaptation, wiring minimisation, noise, neural representation, small-world networks

Säilytyspaikka - Förvaringsställe - Where deposited

Library of Behavioural Sciences, University of Helsinki

Muita tietoja - Övriga uppgifter - Additional information

HELSINGIN YLIOPISTO - HELSINGFORS UNIVERSITET - UNIVERSITY OF HELSINKI

Tiedekunta - Fakultet - Faculty	Laitos - Institution – Department
Käyttäytymistieteellinen tiedekunta	Käyttäytymistieteiden laitos

Tekijä - Författare - Author
Antti Knuutila

Työn nimi - Arbetets titel - Title
Tiedonkäsittelyn tehokkuus aivoissa

Oppiaine - Läroämne - Subject
Kogniotiotiede

Työn laji ja ohjaaja(t) - Arbetets art och handledare – Level and instructor	Aika - Datum - Month and year	Sivumäärä - Sidoantal - Number of pages
Pro gradu -tutkielma	Elokuu 2012	81

Tiivistelmä - Referat - Abstract

Aivojen tiedonkäsittely on huomattavan tehokasta vallitsevien fysikaalisten rajoitteiden puitteissa, jotka liittyvät muun muassa tilan ja metabolisen energian käyttöön keskushermostossa. Luonnonvalinta pyrkii optimoimaan aivoissa toimivien rakenteellisten ja toiminnallisten verkostojen toiminnan yksinkertaisten sääntöjen pohjalta. Nämä säännöt ovat monesti huomattavan samankaltaisia kuin ihmisen suunnittelemissa elektronisissa laitteissa ja tietoverkoissa. Tämä tutkielma esittelee joukon aivojen tiedonkäsittelyn evolutiivista historiaa, kehitystä ja toimintaa ohjaavia yleisiä suunnittelun periaatteita sekä niihin liittyviä tutkimustuloksia.

Keskushermoston rakenteeseen laajalti vaikuttava kaapeloinnin minimoinnin periaate on eräs tällainen sääntö, joka vaikuttaa sekä yksittäisten hermosolujen että kokonaisten aivoalueiden rakenteeseen ja sijoitteluun. Yksittäisten hermosolujen pienentäminen säästää tilaa, mutta vaikeuttaa viestintää kasvattamalla satunnaisen aktiviteetin eli kohinan määrää hermosoluissa. Sekä aivojen rakenteellisissa että toiminnallisissa verkostoissa havaitaan monin niin kutsuttu pieni maailma –tyyppinen rakenne, joka tuottaa tehokkaan verkostorakenteen verrattain pienillä biologisilla kustannuksilla. Hermoviestinnän metaboliset kustannukset puolestaan rajoittavat hermoimpulssien määrää ja luovat evolutiivisen paineen muodostaa tehokkaita neuraalisia representaatioita.

Esitelty kirjallisuus tarjoaa runsaasti todisteita aivojen tiedonkäsittelyn rakenteellisesta ja toiminnallisesta tehokkuudesta. Lopuksi käsitellään nykyisten tutkimusmetodien rajoituksia ja avoimeksi jääviä kysymyksiä.

Avainsanat – Nyckelord - Keywords
optimointi, tehokkuus, adaptaatio, kaapeloinnin minimointi, kohina, neuraaliset representaatiot, pieni maailma -verkostot

Säilytyspaikka - Förvaringsställe - Where deposited
Helsingin yliopiston kirjasto, Keskustakampuksen kirjasto, Käyttäytymistieteet / Minerva

Muita tietoja - Övriga uppgifter - Additional information

“In nature’s infinite book of secrecy
A little I can read.”

– William Shakespeare
The Tragedy of Antony and Cleopatra

Acknowledgments

Mom and Dad, for your infinite patience on which I have drawn heavily.

Otto Lappi and Christina Krause, for their helpful comments.

Kai Kaila, for inspiration in choosing a topic of study.

Helsinki, 21st August 2012

Antti Knuutila

Table of contents

List of figures

List of tables

1 Introduction

The human information processing apparatus, consisting of the nervous system and culminating in the brain, is widely hailed as a distinctive feature of man separating him from other animals. A large brain provides considerable utility in the form of a mental and behavioural flexibility that seems unattainable in its absence. This utility, however, does not come cheap, as brains are extraordinarily complicated structures with exceedingly high operating costs. Given the considerable rewards associated with a large brain running efficiently, natural selection is thought to have optimised both structural and functional aspects of the brain's information processing to a very high degree.

Over long periods of time, natural selection acting on genotypic variation begets biological adaptation in populations. The essence of the theory of natural selection has been described as "a statistical bias in the relative rates of survival of alternatives", i.e., of genes and the individual organisms carrying those genes (Williams, 1966). The present thesis will concern itself with the evidence, consequences, and implications of neurobiological adaptation and optimisation of information processing in the human brain and that of other species.

The evolutionary push towards ever larger brains allowing for more complex behaviours inevitably collides with constraints set by the limited supply of physical resources such as building materials, time, space, and metabolic energy (Laughlin & Sejnowski, 2003). This perpetual conflict creates a need for efficient biological designs that optimise an organism's fitness, reflecting a need to perform as little work and to employ as few scarce resources as possible to achieve a given task. For the brain as a whole, the one clearly defined task is the most general one in biology: differential reproductive success. With respect to smaller units in the brain, their respective tasks range from the minimisation of

individual neuronal components without sacrificing adequate functionality to finding energetically efficient neural representations for biologically salient information.

Neuronal networks are seen at multiple levels of organisation to follow simple design rules linked to constraint satisfaction and the allocation and minimisation of the use of physical resources. The present study will focus on neurobiological efficiency to do with the use of space and energy in particular. Identifying and eventually quantifying such optimisations of brain structure and function can yield strong evidence of their evolutionary significance and guide further research into their role in human information processing.

Design principles relating to efficiency and parsimony in the structural and functional workings of the brain are found at times to be strikingly similar to those employed in man-made electronic devices and networks. Minimisation of the neural wiring connecting brain regions, for example, faces challenges reminiscent of those encountered in the design of large-scale digital microchips, while limits to the brain's energy supply have brought forth adaptations in its neural signalling network that bring to mind schemes employed in modern technological communications systems.

Like so many branches of science, neuroscience progressed at first largely as an exercise in description. The study of biological optimisation in the brain offers one possible avenue forward, illuminating a host of general design principles directing the evolution of the nervous system and affecting its function at all levels of operation. Perhaps in time this approach can provide a source for a "generative grammar of the nervous system" (Cherniak, 1995), some general rules of efficient design that compactly characterise aspects of the nervous system's anatomy and physiology. As the proverb goes, instead of examining only the trees this approach opens up the possibility of adopting a wider perspective and asking how the forest itself came to be.

The main body of this study is divided into five parts. The introductory part gives a summary of the premises and historical development of adaptationist thinking, addresses some criticisms levelled against this school of thought, and introduces the application of mathematical optimisation theory to questions of biological adaptation. The second part describes possible avenues of neurobiological adaptation to do with the processing and transmission of information in the nervous system, and gives examples of analogies between neurobiological and man-made information systems.

The next two sections list evidence of biological optimisation to do with neural information processing. The first part concerns optimisations to do with the conservation of space in individual neuronal components and larger assemblies thereof, while the latter traces out the consequences of energetic limits on the functioning of the brain's signalling network. The thesis ends with a summation of findings and a discussion of implications for future research.

2 On adaptation

Those organisms better adapted to the prevailing environment produce more offspring on average than do contemporary organisms. Given heritability of a trait, differential birth rates between organisms possessing a trait that is beneficial and those without it eventually allow said trait to grow more common within a population. For instance a prey animal with protective colouring will be more difficult to spot for a predator and hence more likely to survive and reproduce. Genetical heritability will eventually spread such a trait over an ever larger portion of the population. Evolutionary change occurs as a consequence of these changes in gene frequencies in a population.

For a given trait to be considered an adaptation it must meet three criteria:

1. the trait has evolved as a variation of an earlier form,
2. the trait is heritable,
3. the trait enhances the fitness or reproductive success of its bearer.

As a corollary of evolutionary adaptation the population attains an ever better fit with its current environment. Over time less useful phenotypes are weeded out in favour of those better attuned to the challenges posed by a given environment.

Wright (1932) formulated the concept of the fitness landscape, a mapping from an organism's genotype or phenotype to its fitness. Natural selection is here portrayed as a force pushing a population of organisms up a proverbial hill, towards ever greater phenotypic fitness. While the fitness or adaptive landscape as a metaphor for the process of adaptation via natural selection has come under repeated criticism (Pigliucci & Kaplan, 2007, ch. 1; Matthen & Ariew, 2002), some in the adaptationist camp have further developed this line of thought by arguing that natural selection acting on adaptive traits

will over time lead to those traits being optimal, that is, better than any imaginable alternative, and even that such optimisation can be expressed mathematically (e.g. Maynard Smith, 1978; Alexander, 1996). This application of optimisation theory, better established in fields such as economics and engineering, to the study of biological adaptation has proved to be a matter of some contention (Alexander, 2001).

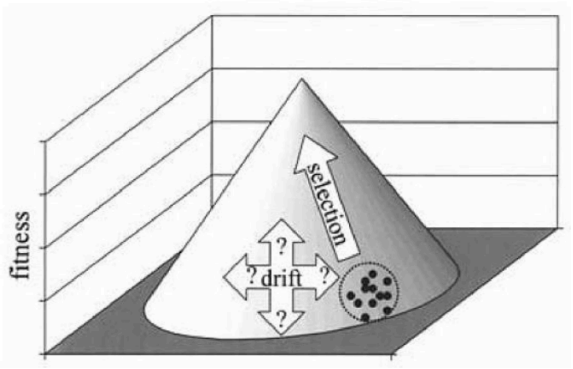

Figure 1. The fitness landscape and the force metaphor. A population is pushed "up" toward peak fitness by the "force" of selection, and in random directions by the "force" of genetic drift. (Pigliucci & Kaplan, 2006, ch. 1)

Any change in environment conditions effectively moves the goalposts of the adaptive game, prompting new adaptations to emerge or old ones to be transformed. Species and their adaptive fitness therefore lag by necessity behind an environment ever in flux and its resultant adaptive landscape. All living things walk an evolutionary tightrope between pressures to adapt and the biological cost of doing so. Evolutionary change is needed in order to adapt to a dynamic environment, whereas constancy and invariance preserve time-tested solutions formed in the course of evolutionary history.

Even a given adaptation's function does not remain constant. Exaptation or preadaptation is a situation where a trait fortuitously takes on a function different from its original one, or where none previously existed (Gould & Vrba, 1982). The wings of alcids, a bird in the auk family, are one example: these birds can "fly" underwater as well as in the air

(Futuyma, 2009, p. 294). Earlier adaptations may also turn detrimental to the organism's survival and reproduction. A trait which is more a hindrance than a help to the organism is known as a maladaptation (Crespi, 2000; Gangestad & Yeo, 1997). Natural selection operates incessantly, but the direction of selective pressures, i.e., which traits are favoured over others and why, depends on the environmental context.

There exist three broad categories of adaptive traits:

1. structural adaptations – those having to do with the physical features of an organism such as the shape or internal organisation of its body,
2. physiological adaptations – those pertaining to biochemical reactions in the organism and involving bodily functions such as hormone secretion or temperature regulation,
3. behavioural adaptations – those concerning what the organism does to survive, e.g., heritable chains of behaviour or a capacity to learn novel behavioural patterns.

As the present study is concerned with the evolution and biophysics of the central nervous system, it will focus mainly on adaptations falling into the first two categories described above. While the organisation of the nervous system has a definite effect on an organism's behaviour, such behavioural adaptation is for the most part beyond the scope of this study.

2.1 Adaptationism

Adaptationism is a theory of evolutionary change that asserts the central importance of selection as a factor in guiding an organism's evolution. Adaptationists contend that natural selection acting on individuals within a population is the only substantial cause for the

evolution of any given trait, and hence the main goal of evolutionary biology is to understand the evolution of adaptations.

Adaptationism as a general notion solidified only in the latter half of the 20th century, but its intellectual lineage can be traced back to the 19th century and the formative years of evolutionary thinking (see Mayr, 1983). The 1930s already saw a lively debate concerning the influence of natural selection and of non-selective forces on the evolution of traits (Fisher, 1930; Wright, 1932). A highly influential adaptationist account was offered by Williams (1966), proposing that natural selection is the only factor that can account for a complex adaptive design's coming into existence.

Godfrey-Smith (2001) identifies three strains of adaptationism. Empirical adaptationism holds that natural selection is the only truly important factor in the evolution of traits, while other evolutionary influences and constraints are of little consequence over the long term. By and large it suffices to focus on selective forces when trying to explain or predict the outcome of an evolutionary process. Under this view, discrepancies between experimental data and the predictions of an optimality model should lead to the model being rejected in favour of a new one.

Explanatory adaptationism, on the other hand, is the view that the "apparent design of organisms, and the relations of adaptedness between organisms and their environments, are the big questions […] answering these questions is the core mission of evolutionary theory" (Godfrey-Smith, 2001). Natural selection is seen as the key to these questions, and is therefore of central importance in evolutionary thinking. In addition to being a claim about evolutionary science, explanatory adaptationism can also be seen as an untestable aesthetic claim more relevant to public discourse on evolutionary biology than science proper.

The third variety of adaptationism has a more practical bent. Methodological adaptationism argues that the track record of selective thinking in biology simply is impressive enough to warrant its continuation. The best way to approach biological entities is "to look for features of adaptations and good design" (Godfrey-Smith, 2001). Adaptation is viewed simply as an organising concept for evolutionary research, and no claim is made about the actual role of selection in the realm of biology.

The three types of adaptationism identified here are thought to be logically independent of each other, i.e., one does not in the strict sense imply either of the two others. However this is not to say that they do not support each other. At a minimum it is useful to note that views regarding empirical and explanatory adaptationism need not go hand in hand.

Lewens (2002) offers a distinction between two kinds of adaptationism. Of these, hypothetical adaptationism is the stronger hypothesis. It outlines the debate over adaptation as an empirical question about the power of natural selection, stating the familiar thesis that selection is the only important cause for most phenotypic traits. The heuristic mode, on the other hand, sees in adaptationism a recipe for generating hypotheses and organising research. Conceived in this way, "adaptationism has no truth value but should instead be judged by its fruitfulness" (Lewens, 2002). This approach is essentially identical to the methodological view of adaptationism described above, and corresponds best with the point of view adopted in this study.

2.2 Critique from panglossianism

An influential critique of adaptationism was offered by Gould & Lewontin (1979), who describe the program of adaptationism as an attempt to "explain the existence and particular forms of nearly any phenotypic trait as the result of natural selection" (Pigliucci

& Kaplan, 2000). In what is colloquially known as the "Spandrels paper", adaptationists stand accused of adopting a panglossian paradigm[1] of baseless optimism regarding the evolution of individual traits in living organisms. Gould & Lewontin (1979) argue that adaptationist thinking tends towards unscientific storytelling and ad-hoc explanations, and should therefore be rejected from serious scientific discourse.

As an architectural byproduct of mounting a dome on top of rounded arches, tapering triangular spaces called spandrels are formed at the intersection of two arches (see Fig. 2). The central dome of the basilica of San Marco in Venice, Italy, features four spandrels, each decorated with mosaic images. While a casual observer might jump to the conclusion that spandrels are purposefully designed as a place to display religious art, their existence is in fact mere architectural happenstance and not a consequence of purposeful architectural design.

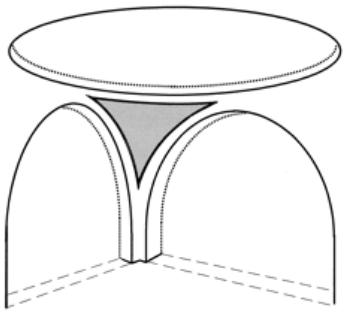

Figure 2. An architectural spandrel. The area marked in gray shows an architectural spandrel between two arches. Image courtesy of the Wikimedia Foundation.

[1] The notion of panglossianism takes its name from Voltaire's *Candide*, in which ridicule is heaped on a certain Dr. Pangloss for believing that "Things cannot be other than they are [...] everything is made for the best purpose."

Gould & Lewontin (1979) wield spandrels[2] as a metaphor to illustrate their critique of adaptationism. The authors argue that adaptationists are all too eager to identify individual traits in organisms and to assume that they are of necessity a result of natural selection (see also Reeve & Sherman, 1993). Standards for adopting and revising hypotheses of how and why a given adaptation came to be are portrayed as too lax to support scientific explanations of trait evolution in living organisms.

All this amounts essentially to a warning against "confusing function with adaptation" (Nielsen, 2009). Seeing a trait put to a specific use should not be taken as evidence that it has evolved for just that particular reason. For most any trait, the critics point out, a fitting story can be constructed of how it came into being, yet falsifying such hypotheses might well prove impossible in a scientific context. Gould & Lewontin (1979) also caution against viewing organisms as mere assemblages of individual traits, optimal or not, arguing that traits cannot and should not be viewed in isolation but as part of a complex and integrated whole.

The Spandrels paper issued a call for a shift towards more sophisticated adaptive analyses that take into account the developmental processes at work in the formation of traits. Gould & Lewontin (1979) offer a pluralistic approach to evolutionary biology, placing an emphasis on historical contingency and the influence of constraints on adaptation. Selection is seen as constrained and often thwarted by a range of developmental factors, with evolutionary outcomes reflective of historical accident as much as adaptation (Godfrey-Smith & Wilkins, 2006). Lewontin (1983) concedes that an emphasis on adaptation was useful as an organising framework early on in the development of biological theory, but regards it as impeding further progress.

---

[2] Although of little practical relevance to any discussion of adaptationism, it has since been pointed out that these architectural forms are in fact properly called pendentives not spandrels (Mark, 1996).

2.3 Constraints to adaptation

An organism's living environment is but one factor affecting the spectrum of possible adaptation. So-called developmental constraints bias the course of evolution, limiting the variability of phenotype production and thus of adaptability at large. Such constraints can be thought of as boundary conditions to evolutionary dynamics which "define the space of the permissible and probable transformations" that living systems can undergo (Schlosser, 2007). Put simply, the way an organism has evolved in the past biases its subsequent evolution, closing off some avenues and opening others.

Numerous ways have been offered of classifying kinds of developmental constraint. Schwenk (1995) groups these constraints into two classes. Class I constraints are those to do with evolutionary channelling, and can be further divided into functional and structural constraints. This type of constraint channels the patterns created and sorted by natural selection, setting boundaries for the path an evolutionary lineage takes through phenotypic space. Class II constraints consist of developmental and genetic constraints. The focus here is on underlying mechanisms that limit the generation of new phenotypes produced but have little effect on lineage-level phenotypic patterns. Class II constraints thus provide "proximate mechanisms and processes independent of natural selection" (Hall, 1999, p. 100) that are further constrained by class I constraints and selective forces.

Gilbert (2010, pg. 744–746) describes three categories of evolutionary constraints:

1. physical constraints – those constraints that arise as a result of physical laws,

2. morphogenetic constraints – construction rules concerning the ways in which an organism's growth and shape can depart from their species-typical development,
3. phyletic constraints – historical restrictions based on the genetic background of an organism's development.

Some phenotypes which are imaginable in principle are ruled out categorically by physical constraints; such constraints are universal in the sense that they hold always and cut off entire avenues of evolutionary variation. For example wheeled appendages are impossible on a vertebrate, since blood cannot circulate to a rotating organ (Gilbert, 2010, pg. 744). Morphogenetic and phyletic constraints, on the other hand, apply only locally in the sense that their effect is dependent on historical context (Maynard Smith, Burian, Kauffman, Alberch, Campbell, Goodwin, Lande, Raup & Wolpert, 1985).

The present study will draw on knowledge of physical constraints to draw comparisons between actual brain functions and their theoretically optimal alternatives. Both universal and local constraints to adaptation are seen to have distinct effects on the organisation and function of the central nervous system.

2.4 Optimality and adaptation

Some evolutionary theorists have come to emphasize the power of natural selection acting on genetic variation to seek out optimal solutions to "problems" encountered by biological systems (Parker & Maynard Smith, 1990; Alexander, 1996). Talk is of problems and solutions in the sense of genotypic and phenotypic variation leading eventually to differences in the survival rate and reproductive success of individual organisms. Over long periods of time this iterative process of refinement is hypothesised to produce

adaptations that are not just very good but in fact optimal. Adaptive optimisation, in this view, is the quantitative expression of the evolutionary process.

Optimisation theory is a branch of mathematics that concerns itself with studying the extremal values of a function, that is, its minima and maxima. Its practical applications are often to do with finding the best alternative from a larger set by minimising costs, maximizing gains, or by finding a workable compromise that satisfies some given constraints. Borrowing analytical tools from optimisation theory is standard fare in various fields of study, but has met with some resistance in the biological sciences. Much criticism of adaptationist thought in general has hinged on the question of whether traits do in fact evolve towards a theoretical optimum (Gould & Lewontin, 1979), although there is no logically necessary connection between optimality modelling and adaptationism (Potochnik, 2009).

An optimisation problem can be described mathematically as follows (Alexander, 1996): given a function $f : A \rightarrow R$ from some set A to the real numbers, find the element x_0 such that

1. $f(x_0) \leq f(x)$ for all x in A (the minimisation case), or
2. $f(x_0) \geq f(x)$ for all x in A (the maximisation case).

This general framework can be applied to any number of real-world problems. Economy and efficiency being “universal characteristics of biological mechanisms” (Williams, 1966, pg. 41), the use of optimisation theory has made it possible for biologists to move “from merely describing patterns or mechanisms to being able to predict, from first principles, how organisms should be designed” (Sutherland, 2005).

Optimisation models relating to biological adaptation contain an assumption about three factors: relevant constraints, an optimisation criterion, and the assumption of heredity

(Maynard Smith, 1978; Sutherland, 2005). Constraints are typically to do with the set of possible phenotypes on which selection can operate, describing the limits of what can evolve in the first place. In the context of this study the relevant constraints are more often related to the use of limited physical resources. The ideal optimisation criterion is an organism's inclusive fitness[3] (Vilarroya, 2002), but since the impact of a single adaptation on fitness is often impossible to judge accurately it is generally sufficient to focus on a specific component of fitness instead, such as the rate of energy intake while foraging (see Fig. 3). Heredity of adaptations is assumed as a precondition to natural selection.

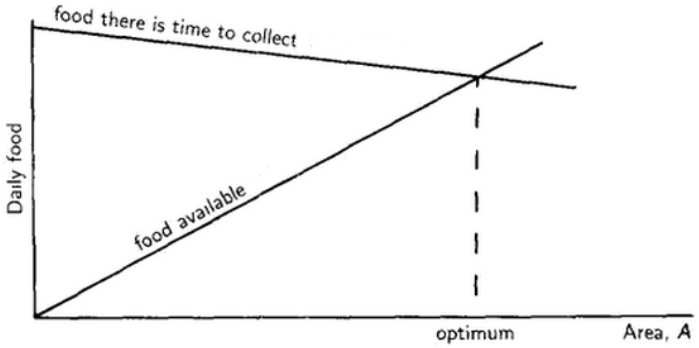

Figure 3. A simplified schematic illustrating optimal territory size for a foraging animal. The quantity of food available and the amount of food there is time to collect are plotted against territory size, demonstrating the effect of constraints on optimal behaviour. (Alexander, 1996)

(Maynard Smith, 1978) further specified that "in testing a model we are testing not the general proposition that nature optimises, but the specific hypotheses about constraints, optimisation criteria, and heredity […] we test whether we have correctly identified the selective forces responsible for the trait in question." This empirical view holds that optimisation methods are a helpful tool in the testing of adaptive hypotheses, and that their usefulness does not rest on the assumption that natural selection optimises always.

[3] Inclusive fitness is defined as the sum of an organism's classical fitness, which consists of its individual reproductive success, and its promotion of the reproduction and survival of genetically related individuals.

Building an optimality model that yields the observed phenotype as an optimal solution is seen as supporting the hypothesis of which factors are at play in the evolution of a trait (Sutherland, 2005; Parker & Maynard Smith, 1990).

2.5 Adaptationism reformed

Since the criticisms raised against adaptationism, most poignantly by Gould & Lewontin (1979), the role played by constraints, tradeoffs, costs and benefits in trait evolution has been a subject of frequent discussion. Constraints to adaptation have come to be seen as "a common and inescapable feature of living systems" (Pigliucci & Kaplan, 2000), with accompanying implications for optimality models. In light of the evident dangers of applying the methodology of adaptationism incorrectly, what possible explanation could there be for its enduring popularity, and has the critique against adaptationist thinking and optimality models been successfully addressed by theoretical advances?

A straightforward response has been to point out the heuristic value of adaptationist explanations (Lewens, 2002). The search for the function of biological structures has been "for centuries the basis for every advance in physiology" (Mayr, 1983), and its continued use is warranted on the basis of this distinguished track record. Whether this line of argument be seen as falling under the heading of methodological (Godfrey-Smith, 2001) or heuristic adaptationism (Lewens, 2002), its adherents can point to past results as evidence that the adaptationist approach is in fact a useful tool in studying evolutionary adaptation. Surely, the thinking goes, adaptive hypotheses could not have yielded such plentiful results if there were something amiss in the underlying theoretical construction.

Parker & Maynard Smith (1990), on the other hand, claim that the continued criticism of the use of optimality models in biology is based on an erroneous understanding of the

purpose for which these models are built. The critique from panglossianism accuses the optimisation approach of clinging to the idea that living organisms are optimally adapted and of seeking to find proof of this, i.e., of erroneously employing a "backwards" mode of reasoning (Pigliucci & Kaplan, 2006, ch.7). However perfect adaptation is not assumed in optimality models at all, but rather the aim is "to understand specific examples of adaptation, in terms of the selective forces and the historical and developmental constraints operating" (Parker & Maynard Smith, 1990).

Since optimality models include, in addition to an optimisation criterion and an assumption about the heredity of traits, an additional assumption about any constraints relevant to a trait's evolution, they should not be held guilty of espousing a panglossian worldview of unqualified optimality (Parker & Maynard Smith, 1990). Crucially, taking into account constraints on adaptation is exactly what precludes perfect adaptation. Where experimental data do not fit the predictions of an optimality model, the original model and its assumptions should be rejected and improved upon. Such an empirical approach sees adaptationism as a general methodology guiding research rather than a specific, testable hypothesis. As such its utility is contingent on the outcome of tests run on optimality models and not something to be decided a priori on theoretical grounds (Orzack & Sober, 1996). Even the failure of a carefully constructed model is instructive in that it tells us that additional factors are at work which are not captured by the model at hand.

A clarification of the relationship between the optimality approach and adaptationism in general was offered by Potochnik (2009) making a distinction between two different uses of optimality models. The strong use amounts to a claim about representation: the model is thought to accurately represent the dynamics of selection, leaving no important evolutionary factors unaccounted for. Such a definition links the strong use of optimality models with adaptationism and the criticism levelled against it.

The weaker claim regarding optimality models is that such a model represents only the role of natural selection in bringing about an evolutionary outcome instead of representing all evolutionary influences (Potochnik, 2009). Perceived in this manner, the legitimacy of optimality models does not turn on that of adaptationism proper. More specifically, the critique raised by Gould & Lewontin (1979) has no bearing on such use of optimality models.

Kangaroo locomotion provides an illustrating example of the influence of developmental constraints on trait optimisation (Potochnik, 2009). Kangaroos travel by a series of leaps, meaning that changes are likely to evolve which improve their ability to move in this particular way. Should bipedal running prove to be an even better mode of locomotion it would nonetheless be unlikely for kangaroos to evolve the ability to run, since initial changes in the transformation from leaping to running would tend to be significantly maladaptive (Maynard Smith et al., 1985). This example also brings to light a more general limitation of optimisation models: given a range of possible trait values, these models can be used to find local optima, not global ones. Although a switch to bipedal running might be the best option in a global, panglossian sense, the phenotypic space of kangaroo trait evolution is constrained in such a way as to make it highly unlikely to occur on anything but the longest of time scales.

What emerges, then, is a "synthesis of constraints (spandrelism) and selection (panglossianism)" (Pigliucci & Kaplan, 2000), where selection works within phenotypic boundaries defined by prevailing constraints. Natural selection and its constraints are identified as the two main players on the evolutionary stage. While it is still more difficult to "test adaptive hypotheses rigorously than it is to make them up" (Pigliucci & Kaplan, 2006, pg. 129), there are unequivocal limits on the extent of trait optimisation accomplished by natural selection. Since new adaptations always build on existing ones,

biological design is about compromising and satisficing (Brownlee, 2007; Simon, 1956) as much as optimising.

2.6 A note on teleology

A notable feature of evolutionary biology in general and adaptationist theories in particular is the recurrent use of terms such as “function” and “design”. References abound to adaptations as having evolved to perform a certain function, to satisfy a need, or to have been designed with a given purpose in mind. On a literal interpretation these terms imply the presence of final causes in nature, an agency akin to that ascribed to human actions. Such parlance is more commonly called teleological, i.e., goal-oriented, and is generally avoided in contemporary science because of its historical association with pre-Darwinian notions of conscious design by a supernatural creator.

A linguistic dilemma thus appears at the very core of the program of evolutionary biology. Put simply, the language employed in adaptive explanations is frequently at odds with the process of evolutionary causation. To talk of natural selection as an optimising process, and of adaptations as solving particular problems that organisms face, is to give the impression of evolution working towards a definite goal. Teleological language is controversial in a scientific context precisely because it seems to presuppose either reverse causation or design for a purpose.

Mayr (2004) addresses the issue of goal-directed language by pointing out that biology studies questions of a different type from those of other natural sciences. In the physical sciences, only questions of “what?” and “how?” are held to be valid. The life sciences, on the other hand, also ask the question of “why?”. Examples are questions such as why an organ is built in a certain way, or why an organism behaves the way it does.

Although evolutionary phenomena are subject to the same universal laws as all other physical phenomena, any explanation of a product of evolution must by necessity take the form of a historical narrative reconstructing its evolutionary past (see Mayr, 1983). Since adaptations do indeed have a purpose in the evolutionary sense, namely the tasks for which they are adaptations, it is the very concept of an adaptation which underwrites the continuing use of teleological language in evolutionary biology (Allen & Bekoff, 1995). In this view, the use of goal-oriented terms in adaptationist explanations is more linguistic artefact than ontological commitment, and as such presents no additional theoretical complications.

Vilarroya (2002) distinguishes between two types of optimality. The evolutionary kind occurs as a result of natural selection favouring the trait which maximises the organism's fitness, whereas in the case of reverse-engineering optimality a mechanism is designed to comply with its intended function. Given a problem to solve, an engineer moves from problem to solution, whereas evolution acts blindly and stumbles upon the solution by chance and brute force. To conflate these two strategies is to open the door to misunderstandings regarding cause and effect.

In biological systems, adaptation and optimisation occur as a result of natural selection acting on the results of random processes, not those of conscious design. Teleological notions are a distinctive feature of the life sciences and frequently helpful as explanatory shorthand, yet one should be mindful of maintaining translatability with strictly naturalistic definitions.

3 Avenues for brain optimisation

The principle of parsimony is frequently invoked as a useful guide in the study of evolutionary biology (Stewart, 1993), and was already noted early on in the study of the nervous system (see Ramon y Cajal, 1995). Other things being equal, natural selection is thought to seek out the least resource-intensive alternative to carry out a given task. Physical resources such as energy, building materials, time and space can be thought of as mediums of exchange that enable the quantitative comparison of competing solutions. These resources are conserved where possible, leading to increasingly efficient biological designs in terms of implementation and maintenance costs.

Raising the question of whether there are physical limits that constrain the human brain's processing power and its growth, Hofman (2001) offers a division of such factors into two categories with two subcategories each:

1. energetic limits to do with blood flow:
    (a) thermal limits – as the brain is actively cooled not simply by heat conduction from the surface of the head but also by blood, a limiting factor to brain growth is how fast heat can be removed from the brain by the flow of blood,
    (b) metabolic limits – the high metabolic cost of the brain requires more blood to be transported to a larger brain;
2. neural processing limits to with the processing of information:
    (a) design limits – the neural architecture needed to process ever growing amounts of information,
    (b) operational limits – the accumulation of signal processing delays in a larger brain.

Hofman (2001) concludes that there is little room for incremental improvement available within the limits of the existing *Bauplan* or construction plan of the human brain. Given the constraints described above and the apparent continued growth of brains over evolutionary time, there comes a point corresponding to a brain volume two or three times that of modern man where the brain reaches its maximum information processing capacity. Once the brain has grown to the point where the bulk of its mass is in the form of neural connections, any further increases in size will be detrimental due to increased signalling delays and declining neuronal integration. Any further evolutionary advances in intelligence would then need to take place outside the nervous system, perhaps in the technological realm with its radically different selection mechanisms and forces.

Humans are seen in this view as being at the cusp of a hard limit to the information-processing capacity of neuron-based systems (Hofman, 2001). It must be noted, however, that even if these conclusions should prove to be correct, growing brains two or three times larger than exist today would likely take millions of years even under the most optimistic of assumptions.

The present study will concern itself with describing adaptive optimisations in the brain to do mainly with four physical resources:

1. space, i.e., the volume taken up by the neural machinery,
2. building materials consumed in the construction and maintenance of said machinery,
3. metabolic energy, consumed mainly by neural signalling in the form of action potentials,
4. the time expended on neural signalling and computation.

Taking into the account the limited space available for this thesis, we will focus in particular on optimisations to do with the use of space and the consumption of metabolic energy. As all four physical resources mentioned above are oftentimes intimately intertwined in the brain, optimisation of one can and frequently does lead to more efficiency also in the use of others. For example, construction of a brain that is smaller in volume will of necessity also require less in the way of biological building blocks. In other cases the need for preservation of one resource is found to stand in direct opposition to parsimony in the use of another.

3.1 Spatial optimisation

Reducing the size of an organ is generally advantageous as long as adequate function is preserved. A smaller brain, for instance, consumes less energy than would a larger one. Shorter distances between brain regions also make for more rapid signalling between different parts of the brain, allowing for quicker responses to external stimuli and more complex information processing within a shorter period of time.

Given the benefits of parsimony in building a neuronal network, natural selection is expected to optimise the brain's structure to allow for efficient performance in as small a volume as possible without sacrificing function. Space-efficient design should thus be visible at all levels of operations, from single neurons to entire brain areas and the neural wiring connecting them.

There are three ways to reduce the size of a nervous system (Laughlin & Sejnowski, 2003):

1. reducing the number of neurons required for adequate function;
2. reducing the average size of neurons;

3. optimising the layout of neurons so as to reduce the lengths of their interconnections.

Since knowledge of the number of neurons required for adequate function of the human information-processing apparatus remains out of reach of contemporary neuroscience, the present chapter will concern itself only with the latter two of these three options.

3.1.1 Miniaturisation

Axons and dendrites connect neurons to each other, relaying signals between nerve cells. They account for the majority of an individual neuron's volume (Goldberg, 2003), hence any minimisation of neuronal volume would be most visible here. Reducing the size of these components allows neurons to be packed closer together, leading to smaller overall brain volume and a decrease in signalling delays within the brain.

Two constraints limit the reduction of axon diameter. Firstly, the molecular machinery required to propagate action potentials, including, e.g., the cell membrane, ion channels and pumps, and a supporting cytoskeleton, can only fit into an axon with a diameter of 0.06 μm or larger (Faisal, White & Laughlin, 2005). This sets a hard limit on the diameter of axons of the type in existence today.

In unmyelinated axons, random fluctuations in the axon membrane potential, also known as ion channel noise, sets the limit at a larger diameter of around 0.08—0.1 μm (Faisal et al., 2005). In axons with a smaller diameter, spontaneous electrical activity precludes the precise timing of action potential generation. As the ratio between signal and noise grows smaller, reliable transmission of information between different parts of the brain becomes ever harder, impairing the brain's normal signalling functions.

Myelination, encasing axons in sheets of electrically insulating plasma membrane, offers another avenue for the miniaturisation of single neurons. While the conduction velocity of action potentials increases with axon diameter, myelination increases the velocity of action potential propagation by a factor of ten or more compared to unmyelinated fibres of the same diameter (Hartline & Colman, 2007). As a consequence, the speed of information processing in the nervous system is increased and reaction times decreased. A similar improvement in conduction velocity achieved through increasing axon diameter would entail an axon on the order of 100 times larger.

Random ion channel noise is ever present in the nervous system on account of the inherent unreliability of its biological components, especially those to do with channel gating in a neuron's dendrites and axons (Sarpeshkar, 1998). The nervous system can, however, also benefit from noise due to the phenomenon of stochastic resonance, a general phenomenon with applications not just in biology but also in physics, chemistry and engineering. Stochastic resonance describes a situation where "the mere addition of random noise to the dynamics improves a system's sensitivity to discriminate weak information-carrying signals" (Hänggi, 2002).

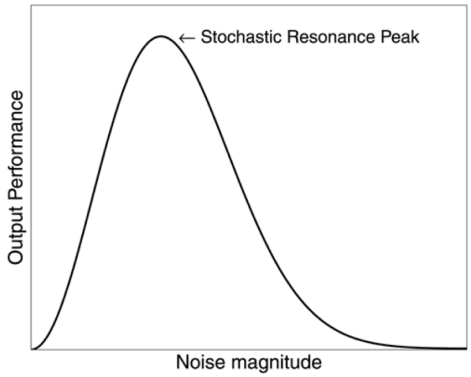

Figure 4. Typical curve of output performance as a function of input noise magnitude in the presence of stochastic resonance. Performance of the system peaks at an intermediate noise level. (McDonnell & Abbot, 2009)

A high level of random noise in the nervous system will deteriorate detection of signals, yet there exists a lower, optimal level of noise which, correlated suitably with a given signal, can aid the chance of detecting sub-threshold signals that would otherwise go unnoticed (Traynelis & Jaramillo, 1998). The nervous system will presumably have adapted to take advantage of the increased signal-to-noise ratio provided by stochastic resonance.

3.1.2 Large-scale minimisation

In addition to miniaturisation of single neurons, space can be utilised more efficiently in the brain by optimising the layout of neurons so that minimal wiring is needed to connect them. The relative frequencies of local (short-range) and global (long-range) connections in particular are a central factor in wiring optimisation, with a greater degree of shorter connections allowing the neural network to be packed into a smaller volume.

The six-layer cerebral cortex is a distinguishing feature of mammalian brains ranging across species from the pygmy shrew all the way to elephants, maintaining a constant thickness of just a few millimeters across species extending over five to six orders of magnitude (Zhang & Sejnowski, 2000). While the size and division of cortical laminae vary greatly between different brain areas and species, mammalian cortex is consistently organised into six main layers (Briggs, 2009). Given that neurons in the human cortex have on average 29 800 synapses for a grand total of some $3.6*10^{14}$ synapses in all of cortex (Roth & Dicke, 2005), finding the one optimal layout for all synaptic connections is practically impossible on any time scale.

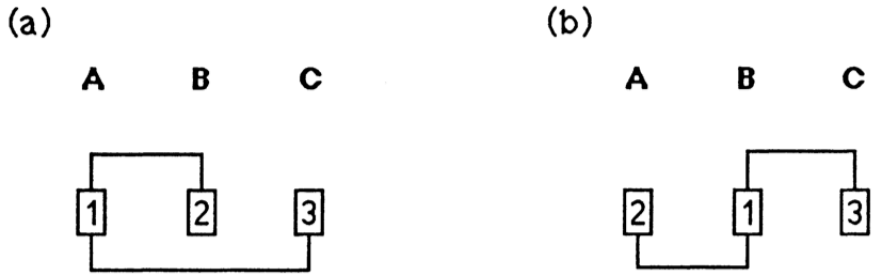

Figure 5. A simple three-component placement optimisation problem. Shown are two alternative placements of three elements in fixed positions. Placement scheme (b) *requires a shorter overall connection length than scheme* (a). *(Cherniak, 1994)*

According to the wiring optimisation principle, among functionally equivalent arrangements of neurons, that is, those arrangements that preserve a given adequate function of the nervous system, "the one having connected neurons as close as possible is most evolutionarily fit and, therefore, most likely to be selected" (Chklovskii & Koulakov, 2004; Raj & Chen, 2011). The biological cost of wiring neurons arises not only from the volume taken up by wiring, but also from factors such as metabolic requirements, signalling delays, and attenuation. This has lead to the idea of optimal component placement in the nervous system, the notion that neural modules of different sizes should be arranged so that any possible rearrangement of these modules would lead to an increase in the total length of wire (Kaiser, 2007; see Fig. 5).

An analogous problem of component-placement optimisation can be found in the design of very large-scale integrated (VLSI) microchips (Cherniak, 1995), and has been shown to be part of the class of NP-hard (non-deterministic polynomial-time hard) mathematical problems. It is conjectured that there exists no solution to this class of problem which can be computed in polynomial time, hence finding an optimal solution to such a problem is prohibitively expensive in a computational sense. In practical applications, NP-hard

problems are typically solved non-optimally using "quick and dirty" heuristic mechanisms instead of exhausting the entire space of admissible solutions.

Finding very good wiring optimisation in the brain raises questions regarding the mechanisms by which such optimisation is accomplished. A simple brute-force search of even a 50-component problem would require timescales longer than the age of the universe[4], even at unrealistic speeds of optimisation (Cherniak, 1994). Given the difficulty of finding an optimal or near-optimal solution to problems of this class, evidence of a high degree of optimisation in the layout of components in the brain would strongly implicate the importance of wiring cost as an organising factor in the brain's evolution.

The central nervous system can be modelled as a giant biological signalling system in which neurons convey information in the form of electrical impulses between different areas of the brain. Such a system can readily be represented as a biological network, an abstract representation of a biological system that suppresses needless detail yet captures its essential characteristics (Alon, 2003). Neurons are here represented as nodes in a network, with their connecting dendrites and axons corresponding to the edges between nodes.

Network analyses based on mathematical graph theory have been widely deployed in fields such as physics and sociology in recent decades. The study of networks is based on the premise that the behaviour of complex systems – be they composed of molecules, neurons or people – is driven by the interactions among their constituent elements (Bullmore & Sporns, 2009). Details of individual interactions between elements of a network are

[4] An exhaustive search of a system with N elements requires searching through N! alternative layouts, therefore an indiscriminate search of all realisable layouts quickly becomes computationally intractable as the system grows in complexity. The computational requirements can in some cases be alleviated by employing suitable heuristics to constrict the search space.

reduced away in favour of a larger picture capturing the essential attributes of the system as a whole.

The lack of a complete connection matrix of the human brain, known as the human connectome[5], presents a formidable challenge to any large-scale analysis of human brain networks. Despite a recent proliferation of neuroscientific databases (Bassett & Gazzaniga, 2011; Sporns, Tononi & Kötter, 2005) coupled with new computational methods under the umbrella term of neuroinformatics (Wood, 2011), a picture of the connectome as a whole still remains out of reach for neuroscience. Even partial datasets can, however, provide glimpses of discoveries to be made once a more complete picture of the human connectome emerges.

Small-world networks are a type of random graph characterised by sparse global connectivity and short average distance between any two nodes, first described in a seminal paper by Watts & Strogatz (1998). This type of network treads the middle ground between the two traditional models of a regular lattice (neighbourhood-connected) network and a completely random one in terms of local clustering and average path length between nodes (see Fig. 6).

Small-world networks are often contrasted with another type of network topology known as the Erdős–Rényi random graph (Erdős & Rényi, 1960), which consists of a network of N nodes so that every pair of nodes is connected with a probability of p. Given small values of p the network is made up of small ensembles without interconnections between them, whereas a larger value of p will produce a continuous network with connections of random length. Such a network has a short average path length with little clustering.

---

[5] The connectome is a deliberate analogy to the concept of the genome, an organism's full complement of genetic information. The Human Connectome Project is working to construct a complete map of structural and functional neural connections in the human brain.

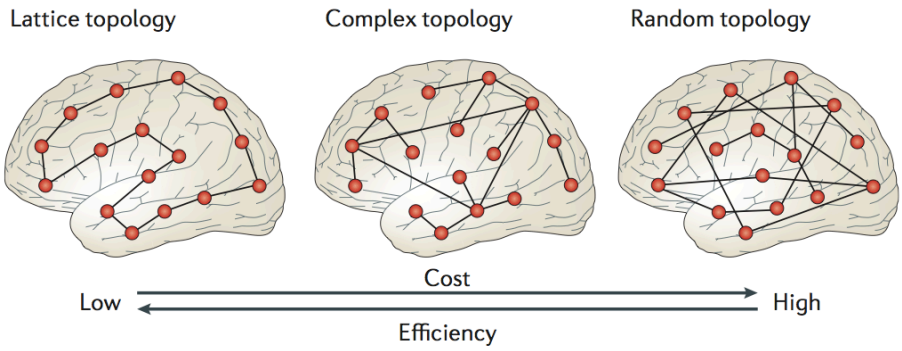

Figure 6. Simplified examples of possible brain network topologies. A regular lattice topology minimises wiring by connecting each node to its two nearest spatial neighbours, sacrificing global efficiency of information processing. A random topology maximises integrative processing of information by connecting nodes at random, yet results in a very high wiring cost owing to the large number of long-range connections. Complex topologies, among them small-world networks, sit between the two extremes, featuring lattice-like local clusters connected by long-distance shortcuts. Such models combine low wiring costs with efficient global transmission of information. (Bullmore & Sporns, 2012)

A regular lattice network, on the other hand, features minimal connection costs but a very long average path length. However in a small-world network neighbours are better connected than in a comparable Erdős–Rényi random network, whereas the average path length remains almost as low as in random networks. A network with small-world characteristics will generally be strongly clustered at the local level, with fairly sparse global connectivity. Meanwhile the average distance between any two nodes, measured in terms of the average number of steps along the shortest path between all possible pairs of nodes in the network, is short. More specifically, in such a network the shortest path L between two nodes chosen at random grows in proportion to the logarithm of the number of nodes N in the network as $L \approx log\ (N)$.

Small-world functional networks come endowed with a number of beneficial features such as resilience against the failure of individual nodes, minimal wiring costs, and efficient signalling between distant parts of the network (Achard & Bullmore, 2007). Given that these are all desirable attributes of a signalling network striving for high efficiency of information transfer at low connection costs, small-world features would be expected to have arisen at multiple levels of organisation in the brain, both in functional and in anatomical networks.

3.2 Energetic limitations

Sending action potentials along neuronal wires is metabolically expensive, accounting for 20% of the energy expenditure of an adult human body and up to 60% for infants (Harris & Attwell, 2012; Laughlin & Sejnowski, 2003). The expensive-tissue hypothesis (Aiello-Wheeler, 1995; Herculano-Houzel, 2012) suggests that the metabolic energy requirements of a large brain are such that natural selection has had to offset them by a corresponding reduction of the gut, "the only one of the metabolically expensive organs in the human body that is markedly small in relation to body size" (Aiello & Wheeler, 1995). Notwithstanding a rise in the organism's basal metabolic rate, a large brain cannot be maintained without reducing the gut by shifting to a higher-quality diet.

Up to 80% of the human brain's energy consumption is related to neural signalling processes (Raichle & Mintun, 2006). Of this portion, another 80% is expended on driving action potentials along axons and creating postsynaptic potentials (Attwell & Laughlin, 2001). The remainder is used, among other things, on maintaining resting potentials and the vegetative functions of neurons and glia cells. The brain as a whole operates under an extreme paucity of energy reserves, with an interruption in blood circulation to the neck of only 7 s leading to loss of consciousness (Ames, 2000).

While the potential bandwidth of the human cortex has been estimated to be as high as a terabit per second (Laughlin & Sejnowski, 2003), energetic constraints dictate that this capacity is never achieved in practice. Energy supply to the human brain sets a limit on the permissible rate of neural signalling of less than 1% of neurons sending action potentials concurrently in the cortical regions of the human brain (Lennie, 2003). Furthermore, the majority of energy consumption is related to intrinsic activity in the brain, with perhaps only 0,5–1% of the total energy burden associated with momentary demands of the environment (Raichle, 2006). The brain can thus be seen to consume energy constantly at rates close to its maximum capacity.

A limited energy supply brings about a multitude of adaptive pressures in the brain. Minimising the amount of neural wiring reduces energy consumption as signalling distances between different areas of the brain are shortened, but miniaturisation of neural components eventually raises problems about noise in signalling (Faisal, 2005). A reduction in size is thus beneficial only as long as adequate function of the signalling network is maintained in the form of a satisfactory signal-to-noise ratio[6]. Neural representations, too, are affected, and must make do with a limited amount of costly action potentials.

3.2.1 Efficient neural representation

In addition to enforcing metabolic efficiency, the high energetic cost of transmitting nerve impulses also implies an efficiency of neural representation. Efficient representation of the incoming signal was already described as a task of the visual system more than fifty years

6 The buildup of noise in sequential stages of neural computation has even been suggested to set a fundamental limit on the logical depth to which brains can compute (von Neumann, 1958, pg. 82).

ago, prompting the suggestion that information theory could provide a quantitative link between the statistical information contained in environmental scenes and neural representations thereof. Barlow (1961) put forth the efficient coding hypothesis, proposing that instead of being equally capable of processing any inputs, the brain and especially its sensory systems should be expected to exploit the statistical dependencies contained in natural stimuli.

Redundancy is a central concept of information theory, being defined as the remainder of the number of bits used in the transmission of a message and the bits of actual information contained therein. Less formally it can be described as those bits wasted in the representation and transmission of a given piece of information. As natural scenes are known to contain considerably redundant information, with perhaps less than half the entropy of random noise (Chandler & Field, 2007), the nervous system is expected to take advantage of these redundancies in encoding the sensory environment by repressing irrelevant data and compressing any information that is deemed worthy of further processing. High redundancy thus allows the sensory system to reduce the problem space of how to represent natural scenes by focussing its efforts on those things which it is likely to encounter in the world. In essence the sensory system is expected to be biased, at both evolutionary, developmental, and behavioural timescales, to efficiently handle the types of inputs it will most probably come across (Simoncelli & Olshausen, 2001).

The ability to make distinctions is one of the fundamental capacities which underlie cognition. In an unbounded environment, natural selection would be expected to maximise the combinatorial or, in the case of neural codes, representational capacity of the brain, i.e., its ability to encode distinct items of information.

One of the fundamental questions about neural coding of representations is whether items of information are coded for by the activity of single neurons or by global activity patterns

of the whole population of nerve cells. There are now both theoretical reasons and experimental evidence suggesting that the brain has adopted a compromise between these two extremes, with the degree of sparseness still largely a matter for future research (Graham & Field, 2006).

	representational capacity	**metabolic cost**	**fault tolerance**
local code	very low	very low	none
sparse code	high	low	high
dense code	very high	very high	very high

Table 1. Properties of fractional coding schemes. (Adapted from Földiák, 2002)

In a neuron population of N binary neurons with Np neurons firing on average, the coding scheme with $p=0.5$ is known to have the largest possible representational capacity of 2^N distinct items (Levy & Baxter, 1996). Yet such a dense distributed coding scheme, also known as a holographic code, is not only difficult to implement but also requires maintaining high average firing rates, which is prohibitively expensive given the brain's limited energy supply. Given that the permissible level of signalling traffic in the brain is in fact remarkably low (Lennie, 2003), maximisation of the brain's representational capacity is constrained by the availability of energy. What arises is an evolutionary pressure to represent information in an energy-efficient manner, with as few nerve impulses as possible[7].

[7] As a historical aside, the branch of mathematics known as information theory (Shannon, 1948) was born of an analogous need to quantify, compress, and transmit electrical signals efficiently during the design of early telecommunications networks in the first part of the 20th century.

Local codes occupy the other extreme of fractional coding schemes (Földiák, 2002). Here each item to be represented is coded for by a single neuron while all others stay silent, and no one neuron takes part in the representation of more than one item[8]. Such a scheme has the benefit of being simple to implement and easy to decode. Its representational capacity, however, is very low, since a network of *N* binary neurons can code for only *N* distinct items. A prohibitively large neural network would be required to meet the representational needs of anything but the most simple of organisms. Extremely local neural codes are also highly susceptible to damage in that the demise of even a single neuron would lead to the loss of whatever representation that neuron coded for.

Sparse coding of neural representations offers a favourable compromise between the two extremes of local codes and distributed ones by combining aspects of metabolic and representational efficiency (Graham & Field, 2006). A sparse coding scheme activates a modest fraction of neurons per item to be encoded, growing the representational capacity of the system exponentially from that of local codes while at the same time being comparatively simple to implement. Additionally, sparse encoding of stimuli necessitates the elimination of statistical redundancy from neural representations. A relatively low average rate of neural activity allows for energy-efficient coding within the limits set by energetic constraints, and sparse codes would thus be expect to feature widely in the brain.

Sparseness of a neural representation is typically characterised in one of two ways (Graham & Field, 2006; Willmore & Tolhurst, 2001). Measures of a single neuron's activity over time are referred to as temporal or lifetime sparseness. Since a population of neurons could appear to exhibit temporal sparseness even with all of them firing together, this parameter alone provides little information regarding an entire neuron population's

[8] Extremely localised coding schemes are occasionally the object of neuroscientific ridicule since they seem to imply the existence of so-called "grandmother cells", neurons responding only to a highly specific stimulus such as the image of one's grandmother (Gross, 2002).

behaviour. Population sparseness, on the other hand, is calculated as an average of individual activity patterns across the set of neurons, giving a measure of average sparseness in the unit. One form of sparseness does not by necessity imply the other.

Recording simultaneous activity from a large population of neurons presents a formidable experimental challenge to any investigation of neural coding. Novel techniques such as optical and multiple electrode recording may eventually yield experimental data on the density of neural coding (Bhalla, 2008), but recording the activity of larger units of neurons is largely impractical as of yet. At present much data regarding the sparseness of neural representations comes from single-cell recordings in the brain's sensory systems in particular (Rolls & Treves, 2011; Olshausen & Field, 2005), where input signals are relatively easy to control. Of note is also the fact that extreme sparseness in neural firing rates represents a quandary for experimental procedures in that such activity might easily be overlooked by an investigator hoping to elicit a response in a very small set of neurons.[9]

A striking feature of physiological recordings from sensory cortical areas is the difficulty of creating stimuli that reliably activate nerve cells. While this applies even at lower levels of sensory processing such as in the primary visual cortex, the problem becomes even more severe at higher levels of processing where neurons respond ever more selectively to inputs. Although the efficient coding hypothesis was originally proposed some fifty years ago, experimental technologies have only recently advanced to the point where it is practical to truly begin testing its predictions.

[9] The apparent sparsity of neural activity in the brain has even brought forth speculation of a "dark matter problem" in neuroscience analogous to the famous conundrum in astrophysics that much of the matter in the universe is undetectable or "dark" (Shoham, O'Connor & Segev, 2006).

4 Efficient use of space in the brain

The parsimonious utilisation of space is a fundamental factor in the biological efficiency of the central nervous system. Adaptations to do with saving space can be seen at multiple levels of neuroanatomy, from components of individual nerve cells and the neural wiring connecting them to global architectural designs combining low costs of implementation with efficient information transfer.

A number of experimental results are presented pertaining to the utilisation of space in the brain and limits to its minimisation by way of biological adaptation. We start at the level of individual neurons and the effects of their miniaturisation on noise and neural signalling, continuing on to a discussion of results surrounding the principle of wiring minimisation. Lastly we examine research concerning the presence of small-world features in brain anatomical and functional networks.

4.1 Local minimisation

Neurons form the basic building blocks of the human information processing apparatus, playing a central part in both neural computation and signalling. Given their sheer number, estimated at 100–200 billion in the human brain (Noctor, Martínez–Cerdeño & Kriegstein, 2007), the physical structure of individual neurons is thought to be a perennial object of optimisation by natural selection so as to minimise the size, transmission times, weight, and energy consumption of the brain.

4.1.1 Axon diameter

An action potential, the basic unit of signalling in cell-to-cell communication in the brain, is transmitted by the concerted actions of a large number of voltage-gated ion channels embedded in a nerve cell's plasma membrane. Noise is an inescapable property of brains operating with frequently unreliable molecular components at the nanometer scale (Faisal, Selen & Wolpert, 2008). Reduction of a neuron's membrane area leads to a concomitant reduction in the number of ion channels carrying the electrical current, resulting in a decline in the signal-to-noise ratio which eventually impairs the performance of neural signalling (Laughlin & Sejnowski, 2003). Wiring minimisation via the miniaturisation of individual neuronal components thus reaches a limit at the point where reliable communications break down.

While individual axons are microscopic in diameter, they may extend to macroscopic lengths and hence account for the majority of a neuron's volume. Adding up the cross-sectional areas of the components of molecular machinery needed to generate action potentials gives a theoretical minimum axon diameter of around 0.06 μm (Faisal et al., 2005). However, computational simulations show that axons with a diameter below about 0.1 μm become practically inoperable because single, randomly opening Na channels will generate action potentials spontaneously and disrupt communication between neurons. As axons grow thinner, neural signalling therefore grows noisier and its capacity to carry information decreases (Faisal & Laughlin, 2007). Axons with a diameter of 0.15–0.3 μm already display a significant amount of spontaneous activity, and below this critical diameter of 0.1 μm the frequency of spontaneous action potentials increases exponentially to the point where the axon's refractory period sets a limit on the firing rate and transmission of information becomes impossible (Faisal et al., 2005).

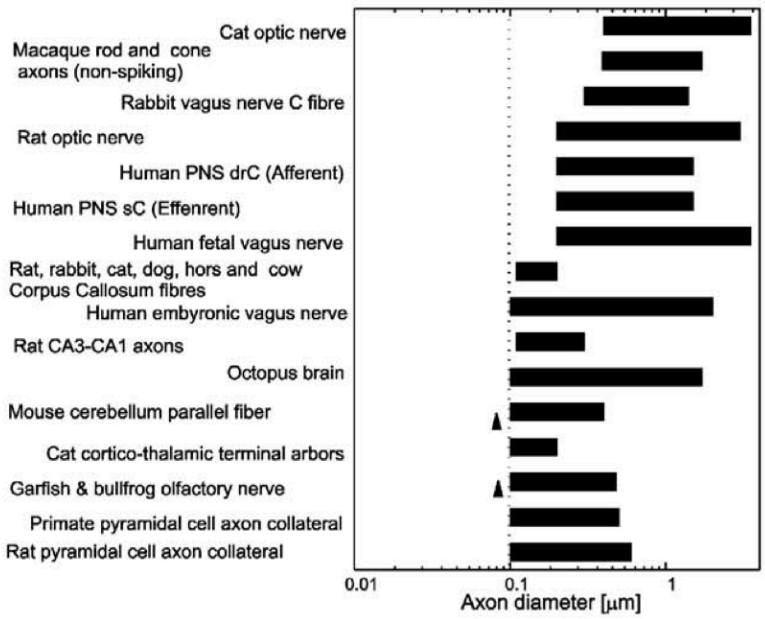

Figure 7. Distribution of axon diameter in a range of species and tissues. The finest axons reach the limiting diameter of 0.1 µm (dotted line); the few exceptions are developing fibers of 0.08 µm diameter (arrowhead). (Faisal et al., 2005)

A review of high-resolution electron-micrograph studies covering several taxa (Faisal et al., 2005) consistently found a minimum axon diameter of 0.1 µm, with rare exceptions down to 0.08 µm. The distributions of axon diameter (see Fig. 7) were found to be skewed towards the critical limit of 0.1 µm, with a typical peak around 0.3–0.5 µm, and to fall off sharply to zero at or just below 0.1 µm. This result is compatible with an evolutionary pressure to minimise axon diameter and suggests that axon diameter is indeed limited at the lower end by increasing ion channel noise.

4.1.2 Myelination

Besides affecting the signal-to-noise ratio of neural communications, axon width is also proportional to conduction velocity (Ikeda & Oka, 2012), advocating an evolutionary

pressure to maximise rather than minimise axon diameter. A biological solution to this apparent contradiction that has arisen independently in several taxa (Hartline & Colman, 2007) is that of a myelin sheath encasing axons in an electrically insulating plasma membrane, allowing for more rapid communications over longer distances without sacrificing wiring minimisation by increasing axon diameter. Myelination also improves the energy efficiency of an axon as it decreases the area of depolarisation during and the amount of repolarisation needed after the action potential has passed (Neishabouri & Faisal, 2011).

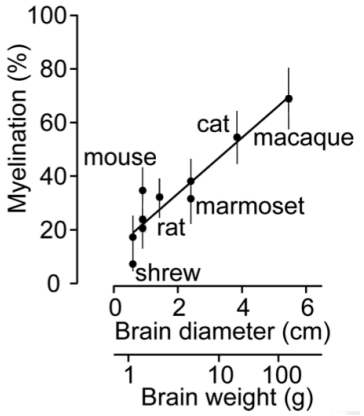

Figure 8. The fraction of myelinated axons increases with brain diameter. Each symbol represents pooled values from one animal. (Adapted from Wang, Shultz, Burish, Harrison, Hof, Towns, Wagers & Wyatt, 2008)

Since organisms are likely selected for fast cross-brain signalling, and nerve impulses will on average travel for longer distances in larger brains, myelination is expected to be more prevalent in larger brains. Wang et al. (2008) report findings in line with this hypothesis, with the ratio of myelinated to non-myelinated axons increasing with brain diameter in species ranging from the shrew and rat to the cat and macaque (see Fig. 8). As a corollary the authors find a corresponding reduction in metabolic cost per action potential created in larger brains, due largely to the higher proportion of myelinated axons. A similar result

concerning the comparative energetic efficiency of myelinated over unmyelinated axons was described by Neishabouri & Faisal (2011).

Evolutionary pressures for minimisation seem to push axon diameter close to its minimum practical width in order to keep down the size of the brain. Myelination helps to increase the conduction velocity of action potentials in larger brains, accommodating faster signalling while still keeping axon diameter close to its practical minimum.

Roth & Dicke (2005) ask the question of how it is that human brains are able to facilitate what is generally taken to be an exceptional mental and behavioural capacity. Either in absolute or relative terms humans do not have the largest brain, the largest cortex nor even the largest prefrontal cortex among different animal species. While the human brain does have the largest number of cortical neurons, elephants and whales are not far behind yet seem to display less intelligence as defined here.

Where the human brain does differ from those of other animals is in its capacity to process information (Roth & Dicke, 2005), and here axonal conduction velocity plays a central role. Myelinated cortical fibres are relatively thick in primates compared to elephants and cetaceans, corresponding to a higher conduction velocity. The authors speculate that coupled with a more tightly packed cortex, i.e., shorter average distances between neurons, this allows for a greater overall information processing capacity and more intelligent behaviour. Additionally, structural and functional specialisations in the human prefrontal cortex are likely to have played an important part in the evolution of human intelligence.

4.2 Noise and stochastic resonance

The probabilistic nature of voltage-gated ion channels, random cell membrane fluctuations, and synaptic processes such as the quantal release of neurotransmitters into the synaptic cleft create electrical channel noise in neural networks (White, Rubinstein & Kay, 2000). As described above, the evolutionary pressure for miniaturisation of neural components makes this noise all the more prevalent. While random noise has the potential to disrupt neuronal communications by impairing the reliability of neuronal responses to given stimuli, noise can in fact play a productive role in the nervous system.

When a neuron is presented with an input that falls below its threshold for firing, it does not produce an action potential and consequently does not transmit any information. If, on the other hand, there is too much random noise coming in from presynaptic neurons, the neuron will fire at a high rate in response to the noise and any outgoing information will be buried under the noise. However the presence of a certain non-zero amount of noise can help trigger the firing of action potentials in response to sub-threshold inputs, thus facilitating the detection of so-called weak signals which would remain undetected in the absence of noise (Kourkoulas-Chondrorizos, 2012; see Fig. 4).

This counterintuitive phenomenon is known as stochastic resonance, and can be described more generally as a situation where "the behaviour of a non-linear system to a weak, periodic stimulus is optimised by the presence of an optimal level of noise" (Traynelis & Jaramillo, 1998). The essential ingredients for stochastic resonance consist of two inputs – a coherent signal and random noise – and an output which is a function of these inputs. Stochastic resonance has been described in a number of physical, biological and technical systems in recent decades (Gammaitoni, Hänggi, Jung & Marchesoni, 2009). With regard to noise in the nervous system, stochastic resonance offers a tantalising hint of non-deterministic processes at work in the brain (McDonnell & Abbot, 2009).

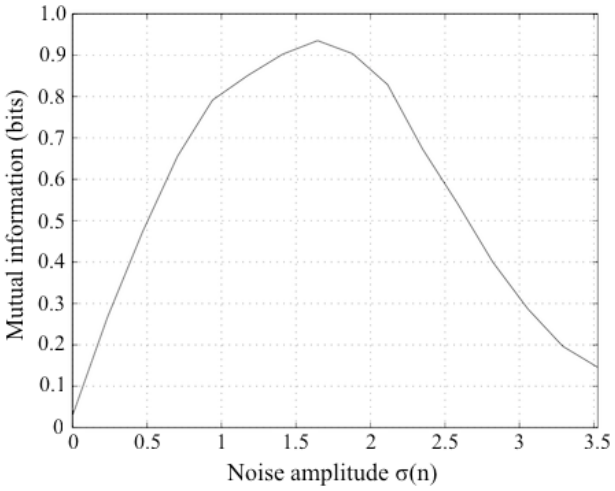

Figure 9. Mutual information as a function of noise amplitude in a population of 100 spiking Izhikevich neurons. Mutual information is defined as the amount of information that is common between input and output signal. The characteristic signature of stochastic resonance is clearly visible (see also Fig. 4). (Adapated from Kourkoulas-Chondrorizos, 2012)

Experimental results provide an abundance of results implicating stochastic resonance in a functional role in neural communications (see McDonnell & Ward, 2011). Phasic neurons in the auditory brain stem have been shown to respond to inputs in the presence of noise which are transparent to them in the absence of it (Gai, Doiron, Kotak & Rinzel, 2009). Rudolph & Dextexhe (2001) describe significantly increased responsiveness to periodic sub-threshold signals in reconstructed neocortical pyramidal neurons simulated according to measured activity in cat parietal cortex. Kourkoulas-Chondrorizos (2012) found noise of a certain amplitude to produce a distinct peak in the mutual information of 100 model Izhikevich[10] neurons, mirroring predictions assuming the presence of stochastic resonance (see Fig. 9).

[10] The Izhikevich neuron model is a simple, semi-empirical model of cortical neurons combining the biological plausibility of the Hodgkin–Huxley model with the computational efficiency of integrate-and-fire neurons (Izhikevich, 2003).

Despite the natural occurrence of noise in the central nervous system, the precise identification of that noise and its incorporation into theories of stochastic resonance remains an unresolved experimental problem (Moss, Ward & Sannita, 2004). While much of the evidence for stochastic resonance at work in the brain is still indirect at this point, i.e., the result of applying both signal and noise externally to neural systems or simulations thereof, it has been argued that "it would be more surprising if the brain did not exploit randomness provided by noise – via stochastic resonance or otherwise – than if it did not" (McDonnell & Abbott, 2009).

Stochastic resonance can be observed in a neural system when a computational goal is better achieved in the presence of noise, that is, random biological fluctuations, than in its absence (McDonnell & Ward, 2011). In the present context the phenomenon of stochastic resonance can be seen as a useful and biologically efficient adaptive compromise between pressures to minimise wiring, to maintain reliable neural communications, and to improve detection of biologically relevant weak signals[11].

4.3 Global minimisation

The wiring economy principle, known also as the wiring optimisation principle, postulates that "for a given wiring diagram, neurons are arranged in an animal to minimize the wiring cost" (Chen, Hall & Chklovskii, 2006). The evolutionary costs flowing from excess wiring are attributed to factors such as the volume taken up by wiring, signalling delays, and attenuation, as well as metabolic expenditures associated with signal propagation.

[11] Additionally, noise is thought to play a part in the development of the brain, increasing with maturation and correlating positively with stable behaviour once it reaches an optimal level (McIntosh, Kovacevic, Lippe, Garrett, Grady & Jirsa, 2010).

Science at present has yet to assemble the human connectome, a complete description of the structural connectivity of the human nervous system (Sporns et al., 2005), precluding whole-brain analysis of wiring cost optimisation in humans. The entire nervous system of some simpler organisms or subdivisions thereof have, however, been mapped exhaustively and provide suitable material for the study of wiring economy (Bullmore & Sporns, 2009).

4.3.1 *Caenorhabditis elegans*

Caenorhabditis elegans is a free-living nematode (roundworm) around 1 mm length. It has been used extensively as a model organism because of its simple structure and transparency (Leung, Williams, Benedetto, Au, Helmcke, Aschner & Meyer, 2008). *C. elegans* is among the simplest organisms with a nervous system, consisting of only 302 neurons in the hermaphrodite (Kosinski & Zaremba, 2007), and therefore provided a self-evident choice for early investigations of component-placement optimisation in the nervous system.

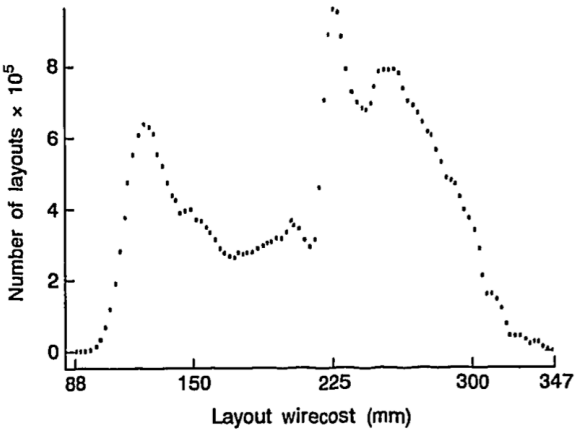

Figure 10. Distribution of neural wiring costs of all 39,916,800 possible layouts of ganglia in C. elegans*. The least and most costly layouts are rarest, with the actual layout corresponding to the smallest wire cost. (Cherniak, 1995)*

Optimal ganglion-level placement in *C. elegans* was reported by Cherniak, (1995)[12] from a set of 11! or 39,916,800 possible layouts, an exceptionally high level of conformance to the rule of *save wire*. The worst possible or "pessimal" layout would correspond to a wiring cost of about four times as much connecting fibre as the actual and optimal layout (see Fig. 10). A brute force search of the relevant placements for all 302 neurons was deemed impractical, as the number of possible permutations increases exponentially with the number of components.

Young & Scannell (1996) offer a critique of the findings reported by Cherniak (1995), pointing out several departures in actual brain connectivity from what is predicted by wiring minimisation. The authors raise the possibility that a simple adjacency rule, a hypothetical tendency of neural components to innervate their neighbours, could account for observed regularities of neuroanatomical connectivity. Minimal connectivity, it is claimed, does not imply component-placement optimisation: "the optimally small length of tunnelling between London Underground stations [...] does not imply that the geographical positions of Piccadilly Circus and Leicester Square have been optimised to minimise the amount of tunnel" (Young & Scannell, 1996).

Cherniak (1996) counters that the wiring optimisation found for ganglia in *C. elegans* can in fact not be achieved by "a computationally cheap procedure such as the adjacency rule [...] when there are too many interconnections for all connected components to be contiguous," nevermind for network architectures typically far more complex than in the nematode. While the layout of the *C. elegans* nervous system does depart strongly from a random pattern in favour of an adjacency rule, such a rule alone does not suffice to produce the actual layout.

---

[12] The paper also included one of the very first complete, single-image depictions of an entire nervous system at synapse-level detail.

The hypothesis that neuronal placement minimises wiring costs as specified by synaptic connectivity was pursued by Chen et al. (2006). For 279 non-pharyngeal neurons in *C. elegans* most were found to be located close to their optimal position, with mean deviations of less than 10% from the optimal layout using different wire-cost models. The authors suggest that wiring optimisation plays an important role in guiding the organisation of the network and, as outliers neurons shared distinct structural characteristics, can even be used to infer neuron function.

More support for the wiring economy of *C. elegans* is offered by Pérez-Escudero & de la Polavieja (2007). The authors dissect the nematode nervous system into smaller sub-networks to test for the existence of smaller, optimised networks within the system at large. The *C. elegans* nervous system was found to contain a subnetwork comprising 84% of neurons wired almost optimally, with a very low position error of 5.4%. A smaller sub-network of sensory and motor neurons containing 15% of all connections can alone predict the main features of encephalisation and the clustering of neurons into ganglia.

Building on the results described above, Pérez-Escudero, Rivera-Alba & de la Polavieja (2009) submit that a simple probabilistic rule can explain the deviations from optimality found in the nervous system of *C. elegans*. All neurons in the nematode with large deviations from their optimal position have a low number of connections with other neurons; in other words, the largest deviations are to be found in those components with the least impact on an indirect measure of fitness derived from the wiring cost. A computational simulation of wiring costs incorporating this consideration shows only 0.033% of all possible permutations producing a smaller wiring cost than the actual layout in *C. elegans* (see Fig. 11).

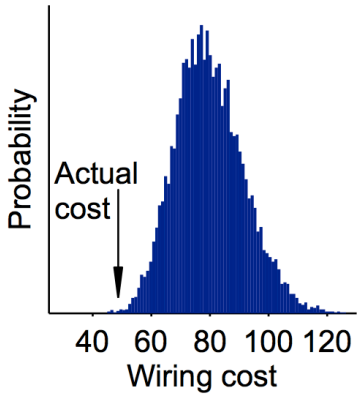

Figure 11. Histogram of wiring costs resulting from random deviations of somatic positions from their optima in C. elegans*. The arrow indicates the cost of the actual configuration, with only 0.033% of configurations resulting in a lower wiring cost. (Pérez-Escudero et al., 2009)*

4.3.2 Separation of white and gray matter

Segregation of the brain into white and gray matter is a ubiquitous feature of vertebrate anatomy. Gray matter is composed mainly of neural somata, synapses, and short-range connections consisting mainly of dendrites and unmyelinated axons. White matter contains long-range connections which implement global communications in the brain, and is largely made up of long, myelinated axons. As this segregation of components is not a necessary condition for achieving the observed connectivity, it is assumed that this particular mode of organisation carries with it some evolutionary advantage.

Wen & Chklovskii (2005) studied the reasons for the segregation of gray and white matter under the assumption that maximal brain functionality requires high interconnectivity of components and minimal conduction delays in the transmission of neural signals. As high connectivity adds to conduction delays because of increased volume and larger distances

between neurons, the two requirements contradict each other, and any compromise between the two must accommodate both their respective needs.

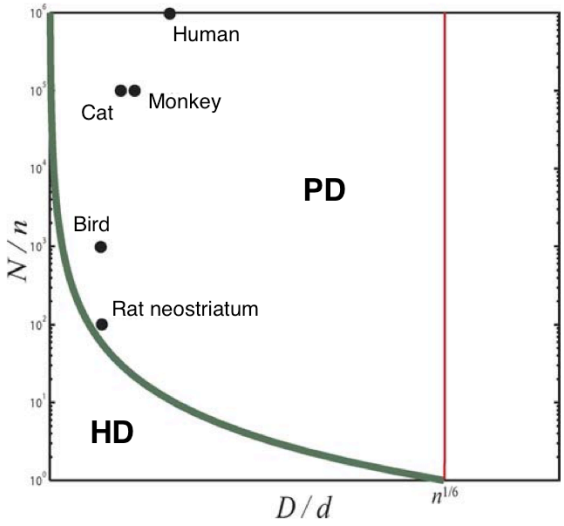

Figure 12. Regimes in which a homogenous design (HD) or a perforated design (PD) separating white and gray matter are optimal. Calculated in terms of the global axon diameter D*, local wire diameter* d*, total number of neurons* N*, and the number of local connections per neuron* n*. Values for the brain of different species are shown. (Adapted from Wen & Chklovskii, 2005)*

Borrowing analytical (non-numerical) tools from theoretical physics, Wen & Chklovskii (2005) were able to show that conduction delays limit the size of a highly connected neural network. In order to avoid signalling delays on the order of seconds, the brain must combine sparse global connectivity with a high degree of local connections. Conduction delays were shown to grow longer in a homogenous design intermixing local and global connections compared to the actual, segregated design (see Fig. 12).

In a related study, Chklovskii et al. (2002) asked the question of what fraction of volume should be taken up by axons and dendrites in optimally wired gray matter. The authors derive an equation describing how this wire fraction depends on four factors: conduction

delay in axons, signal attenuation in dendrites, the number of synapses, and a layout parameter specifying the length of wire used for a given arrangement of neural matter. Varying the wire fraction while keeping all but one of the variables fixed, each of these quantities is found to reach an advantageous extremum when the wire fraction has a value of 3/5. Measurements from three cortical regions in mouse give a wire fraction not significantly different from 3/5 (see Fig. 13), leading the authors to conclude that cortical circuits are organised close to optimal efficiency in the sense of minimising conduction delays, cable attenuation, and total wiring while at the same maximising the number of synaptic connections.

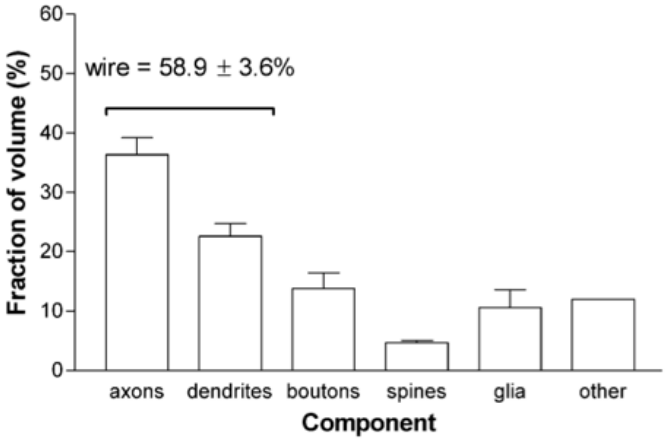

Figure 13. Wire fraction in mouse neocortex. Wiring is found to occupy a portion close to its optimum of 3/5 of gray matter. (Chklovskii, Schikorski & Stevens, 2002)

The anatomical segregation between local and global connections embodied in the separation of gray and white matter also enables investigation of the scaling relationship between the two. Larger brains by definition require longer connections to allow for communication between distant areas of the brain, resulting in disproportionately more white matter as brain grow in size. Zhang & Sejnowski (2000) report a power law with an exponent of 4/3 relating the volume of white and gray matter in species spanning five to

six orders of magnitude. Herculano-Houzel, Mota, Wong & Kaas (2010) reach a similar result, noting that white matter mass scales in primate brains as the power function of the number of neurons in gray matter to an exponent of 1.197 ± 0.091. Larger brains therefore seem to be burdened with the cost of supporting an outsize communications infrastructure in white matter, which grows in size significantly faster than does the amount of gray matter.

4.3.3 Layout of brain areas

Optimisation of neural component placement affects the organisation of the nervous system at multiple levels of hierarchy, and has been the object of extensive study in recent years (see Bullmore & Sporns, 2012; Cherniak, 2012). Cortical areas are brain modules defined by their structural architecture and functional specialisation. In the cerebral cortex, the resource constraint of wiring minimisation actualised through optimisation of the layout of cortical areas has been lightheartedly described as a kind of "plate tectonics of the cortex" (Cherniak, 1994).

Cherniak, Mokhtarzada, Rodriguez-Esteban & Changizi (2004) analysed the global optimisation in the layout of such mammalian cerebral areas where adequate information on connection data is available. In order to avoid the experimental and computational difficulties of measuring actual wire length in the cortex, the authors derive a simpler adjacency cost and a size law that predicts optimisation patterns in sub-networks of cortical areas. For 17 continuous macaque visual cortex areas and 14 cat cortex modules spanning 40 Brodmann areas of visual, auditory, and somatosensory regions, the authors find the actual layouts to correspond to an optimisation of wiring cost in the top 10^{-7} of all possible layouts. Optimality of the layout was found to improve exponentially with sample size,

with larger subsets of cortical areas performing progressively better in their relative ranking for adjacency rule optimisation (see Fig. 14).

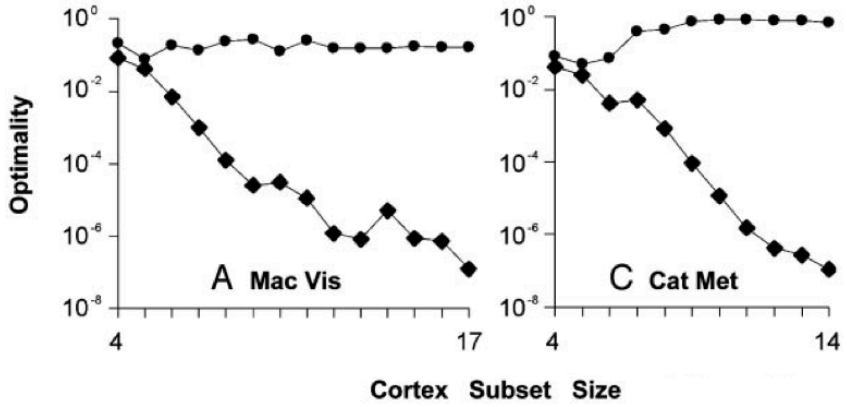

Figure 14. Size law for 17 macaque visual cortical areas and 14 metamodules in cat cortex. Subsets of cortical areas are compared with 10^9 alternative layouts for wiring optimality (conformance to an adjacency rule). The actual layout curve (diamonds) shows smaller subsets ranking toward the middle of their group of alternative layouts, with optimality-ranking improving as subset size increases. For comparison, each scrambled layout curve (circles) shows the corresponding analysis for layouts of the areas with their adjacencies randomly shuffled. E.g., for macaque, fewer than one in 10^6 of all alternative layouts conform to the adjacency rule better than the actual layout of the complete set of 17 cortical areas. (Cherniak et al., 2004)

A study of the layout optimisation of 11 macaque prefrontal cortical areas was carried out by Klyachko & Stevens (2003). The authors compared the total axon volume corresponding to many thousands of alternative arrangements of areas belonging to the orbital-medial network of the prefrontal cortex to the volume corresponding to the actual layout. As in the case of *C. elegans* ganglia described above (Cherniak, 1995), this experimental setup yields 11! or some 39.9 million potential arrangements. The actual layout of macaque cortical areas was found to be wired optimally, minimising the axonal volume required to satisfy existing connections. An additional test was performed by

perturbing the locations of cortical regions in order to verify that the wiring optimisation in evidence was not merely the consequence of an adjacency rule specifying a preference for connections between neighbouring areas. The actual arrangement of cortical areas was found to be optimal for perturbations of inter-areal distances of up to 15%.

4.4 The small world of the brain

Evidence of small-world architecture in functional networks of the primate brain was first described by Stephan, Hilgetag, Burns, O'Neill, Young & Kötter (2000). The authors collated published data on the spread of neural activity in the macaque cerebral cortex in vivo after induced disinhibition. Three independent methods of computational analysis all confirmed that the network of interactions within the cerebral cortex is not homogenous but highly organised into clusters. Cortical areas were found to form three main clusters combining high intrinsic and low extrinsic interactivity: a somatomotor cluster, a visual cluster, and an orbito-insular-temporal cluster. These clusters were also judged to be broadly consistent with known anatomical divisions in the cortex.

A companion paper by Hilgetag, Burns, O'Neill, Scannell & Young (2000) reported similar results for the anatomical connectivity of macaque and cat cortex. Two different computational methods were used to analyse compilations of cortico-cortical connection data in the two species. Macaque and cat cortices were found to display the characteristically small-world features of being organised into densely intraconnected clusters composed of anatomically identifiable member areas, with average local connectivity within the clusters nearly twice as high as the level of interconnectedness between cortical areas. Average path lengths calculated from the connectivity data were on the same order of size as the short characteristic path lengths typical of randomly redistributed data. In the primate visual system, the results confirmed a sub-division into

three main groups of primary, ventral, and dorsal areas similar to those reported by Stephan et al. (2000).

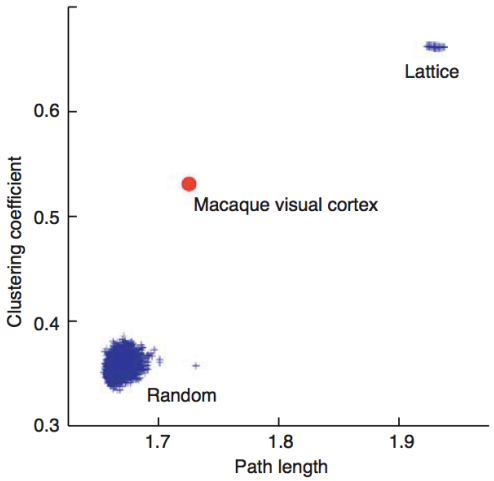

Figure 15. Path length and clustering coefficient[13] for a large-scale connection matrix of the macaque visual cortex. The cortical matrix has a path length similar to a random network, but a much larger clustering coefficient. (Sporns, Chialvo, Kaiser & Hilgetag, 2004)

The results of both companion papers presented above were confirmed and extended in a later study by Sporns & Zwi (2004), who also note the striking tendency of cortical connectivity to display small-world characteristics across multiple spatial scales ranging from large-scale inter-regional pathways to local circuits. Densely connected and functionally related modules are recombined at different scales to produce larger functional and anatomical units in a manner reminiscent of other types of biological systems such as metabolic and protein networks.

Kaiser & Hilgetag (2004) investigated the development of cortical networks by studying the perennial favourites, cat and macaque cortex, additionally employing a simulation of

[13] The clustering coefficient is a measure of the degree to which nodes in a network tend to group together.

neural growth using a simple spatial growth model. Networks resulting from the simulation were compared to actual brain anatomical networks for measures relating to the degree of clustering and the average shortest path between nodes. For both cortical as well as simulated networks the total wiring length was found to sit between those corresponding to minimal and random wiring, indicating a general preference for short-length connections.

Comparing the average shortest path for different wiring setups, longer average path lengths were found for minimal wiring than for the actual networks (Kaiser & Hilgetag, 2004). The authors conclude that the existence of long-range connections in these networks increases total wiring length above its theoretical minimum, but leads also to a shorter average path length. The conservation of shorter paths, that is, a lower number of interconnections in cortical pathways, might thus be a more important constraint in the development of brain networks than achieving minimal total wire length.

Achard & Bullmore (2007) investigated the efficiency and biological cost of functional networks in the human brain using functional magnetic resonance imaging (fMRI) data from two groups of healthy adults. Efficiency was specified in terms of the minimum path length between different brain regions, with costs defined as the total number of edges or connections in the network. The functional connectivity of 90 cortical and sub-cortical regions was estimated using wavelet correction analysis of fMRI time series and then thresholded to produce a set of undirected graphs representing the brain functional network of each participant in the study. The topology of all functional networks was found to be between a random and a lattice graph in terms of local and global efficiency as a function of their implementation costs, a characteristically small-world behaviour. Additionally, the performance of small-world networks was found to decrease with age, with measures of efficiency declining markedly in a control group of older subjects. The authors conclude that functional networks in the human brain have small-world topological properties that

would be expected on the hypothesis of selection for high efficiency of parallel transmission of information at low cost.

Furthermore, the efficiency of both brain structural and functional organisation has been linked to individual differences in human intelligence as measured by the intelligence quotient (IQ). Li, Liu, Li, Qin, Li, Yu & Jiang (2009) calculated topological small-world properties of the brain anatomical network of 79 adult human subjects divided into general and high intelligence groups based on IQ test scores. The authors found significant correlation between network properties and intelligence, with higher IQ scores correlating positively with both a shorter average path length in the anatomical network and higher global efficiency defined in terms of the parallelism of information transfer in the network.

Similar results were reported by van den Heuvel, Stam, Kahn & Hulshoff Pol (2009), who studied the correlation between the topology of brain functional networks and IQ scores in 19 subjects. Topological properties of the functional brain networks were calculated by applying graph analysis to fMRI time series data recorded from the subjects. Again, the authors found a strong positive correlation between the level of global communication efficiency and intellectual performance in the form of a strongly negative association between characteristic path length and IQ.

Interestingly, the number of edges (connections between neurons) was found to correlate significantly with intelligence for brain anatomical networks (Li et al., 2009) but not for functional networks (van den Heuvel et al., 2009). While the brain's structural and functional networks are certainly aligned to some degree, these results present one practical example of how the exact relationship between the two remains elusive.

5 Energy efficiency in the brain

Neural signalling is metabolically expensive to the point of being restricted by the energy supply of the brain. Nerve impulses are processed and transmitted by individual neurons in a manner adapted to this scarcity. Even the extent of neural signalling in the brain as a whole is highly constricted by energetic margins, leaving large swathes of the human cortex silent for the better part of time.

The representation of information in neuronal networks, too, demonstrates energy-efficient behaviour on multiple fronts. Owing to the considerable energetic demands of creating and relaying action potentials, neural representations are composed economically utilising a minimum of costly spikes.

5.1 Signal processing in neurons

Neural networks have evolved in an environment that is inherently noisy, and neurons themselves are relatively unreliable units of computation. Noise is the fundamental enemy of communications engineers (McDonnell & Abbott, 2009), whose goal is to design systems to transmit signals from one place to another efficiently and with as few errors as possible. Signal processing in the brain thus faces an environment that is hostile to its fundamental purpose.

Analogue computational systems operate with continuous values of physical variables in continuous time, while digital system use discrete values and discrete time. Owing to their physical makeup, analogue systems are prone to accumulate noise during sequential stages

of computation, leading to offset errors as computations grow more complex (Sarpeshkar, 1998). Digital systems, on the other hand, restore the value of the signal to one of two attractor states – 0 and 1 – at each stage of computation. Although this error correction ensures the robustness even of complex digital systems, it comes at the price of additional energy expenditure required after every step of computation. At low precision, digital systems are comparatively less energy-efficient as much of their ubiquitous error correction could be sacrificed without substantially affecting the accuracy of computations (Mandal, 2009).

As sensory data are by definition analog, it is beneficial to have at least an analogue front-end to capture their richness. The rest of signal processing in the brain, however, cannot be entirely analogue because noise will accumulate during successive processing stages and eventually degrade the reliability of the computation (Sarpeshkar & O'Halloran, 2002).

Nerve impulses typically travel along dendrites as graded potentials, i.e., analogue signals. These signals are summed up in the cell body or soma of the neuron and, once a triggering threshold is exceeded, an action potential is generated that propagates along the axon. Since the action potential is generated on an "all-or-nothing" basis, meaning that its amplitude is independent of the amount of current that produced it, it is in essence a digital signal. Nerve cells can thus be seen as combining beneficial features of the analogue and digital regimes to produce an energy-efficient hybrid model of computation using comparatively unreliable components (Sarpeshkar & O'Halloran, 2002).

A neuron essentially functions as an A/D or analog-to-digital converter succeeded by its opposite number, a D/A or digital-to-analog converter (Sarpeshkar, 1998). Such an A/D/A device takes as its input a continuous quantity in the form of an analogue signal, then converts it into a discrete representation in digital form, and lastly performs the reverse operation to produce a fresh analogue signal. Carrying signals in analogue form curtails

energy consumption, keeping down communication costs, while the digitisation phase provides error correction to allow for the construction of complex computational sequences.

Besides minimisation of wiring and increased delivery speeds, myelinated axons also work to conserve energy in neuronal activities. In axons covered in a myelin sheath, a nerve impulse's passing results in a smaller ionic imbalance in need of restoration, resulting in a several-hundred fold improvement in metabolic energy efficiency of neural signalling traffic (Hartline & Colman, 2007).

5.2 Energetic limits to neural signalling

Electrophysiological recordings from human cortex show that neurons are intensely activated when engaged in appropriate tasks. The metabolic cost of a single neural spike is high, which severely limits the amount of neurons that can be substantially active at any one time. Signalling-related expenditures account for a large fraction of the brain's energy consumption, which favours the use of energy-efficient neural codes and wiring patterns.

The cost of spiking neurons arises mainly from restoring the ionic imbalances caused by action potentials and from the release and re-uptake of synaptic neurotransmitters. Attwell & Laughlin (2001) analysed the energy expenditure of different neural components on excitatory signalling in the grey matter of rodent brain, concluding that the adenosine triphosphate (ATP)[14] use triggered by each action potential is on the order of 1.6 times greater than usage per second on the resting potential. The total rate of energy consumption

[14] ATP is a multifunctional coenzyme often referred to as the molecular unit of currency used by cells.

thus rises quickly with spike rate, advocating the use of distributed neural codes to improve the energy efficiency of neural representations[15] (see Fig. 16).

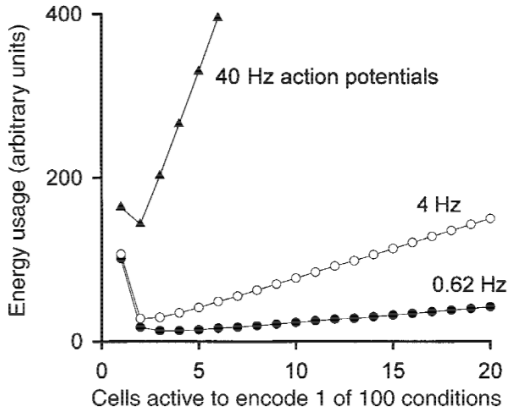

Figure 16. Energy usage in encoding 1 of 100 different conditions as a function of the number of cells simultaneously active. Distributed coding can be seen to improve energy efficiency. (Attwell & Laughlin, 2001)

Scaling up from figures given for rat cortical neurons by Attwell & Laughlin (2001), Lennie (2003) calculated an estimate of the energy consumption of pyramidal neurons in the human cortex. An individual spike in the human cortex was found to be 3.3 times as costly as in rat cortex, while maintaining resting potentials is 2.6 times more expensive. Using these numbers, neocortex is estimated to account for 44% of the human brain's overall consumption of some 20% of all energy consumed by the adult human body.

Given the high energetic cost of neural spiking, only a very small fraction of neurons can be active at any one time. Lennie (2003) gives an estimate of a sustainable spike rate in human cortex of only 0.16 spikes/s/neuron. An average spike rate of 1.8 spikes/s/neuron would expend more energy than is consumed by the whole brain, while a spike rate of around 13 spikes/s/neuron would use more energy than is consumed by the whole body.

[15] The unreliable nature of neural signalling is another factor weighing in on the side of distributed coding: given unpredictable wires, single-channel transmission of information is prone to miscommunication and data loss (Sarpeshkar, 1998).

Even on more conservative estimates of the energy expended on signalling, an average spike rate in human cortex of 3.1 spikes/s/neuron would exceed the amount of energy consumed by the entire brain. Hence with even a few percent of neurons highly active over all of cortex the metabolic burden associated with neural signalling would grow unsustainable. Even with lower average rates of neural activity, there is a need for machinery to allocate resources efficiently across cortical regions according to demand.

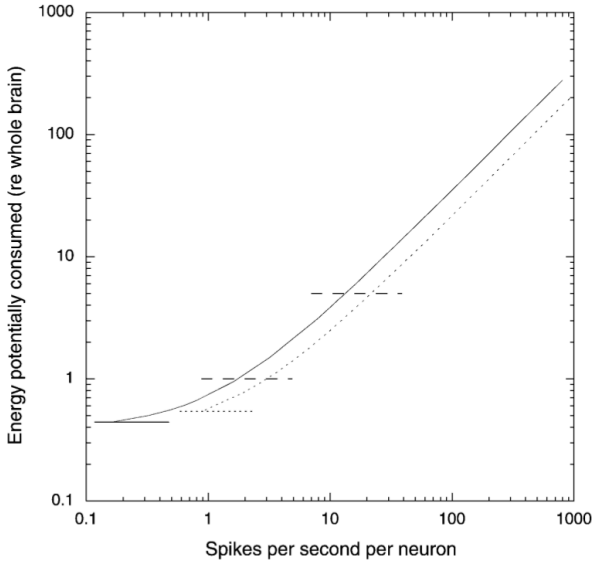

Figure 17. Energy potentially consumed by human cortex as a function of average sustained activity in all neurons, expressed as a fraction of whole brain consumption. The solid curve is based on best estimates of energy supply and costs, while the dotted line corresponds to conservative estimates of energy consumption. (Lennie, 2003)

Attwell & Gibb (2005) studied the effect of energetic limits on the brain's "clock speed", that is, its speed of information processing. Rapid processing of information requires both swift processing of sub-threshold presynaptic potentials and the propagation of action potentials at high frequencies to postsynaptic neurons, which together account for the majority of energy expended on signalling. Diversion of energy to the processing of synaptic potentials would bring on a decrease in the rate of action potential generation and

vice versa. For information processing to occur significantly faster, heavy investments would be needed to, e.g., increase the density of capillaries in the brain. Yet capillaries are already on average only 60 μm apart in the gray matter of the human brain, and a denser capillary network would significantly reduce the space available for neurons in the absence of a corresponding increase in brain size. The energy available for neural signalling therefore seems to limit its timescale of operation to the current millisecond range.

An intriguing connection can also be made between the energetic limits of cortex and the traditional view of attention as an adaptation to the brain's limited capacity for processing information. The scarcity of permissible aggregate neural activity does not imply that only a few tasks can be undertaken concurrently, but it does provide "a natural metric for characterising aggregate task difficulty: how much cortex is active" (Lennie, 2003).

Since local neural activity varies strongly with task while the general rate of energy consumption in the brain is essentially constant (Raichle, 2006), mechanisms are necessary for the efficient allocation of energy. Local changes in cerebral hemodynamics, exploited in functional magnetic resonance imaging, are a clear sign of this machinery in action (Lennie, 2003).

The fact that the brain consumes energy at a relatively constant rate, varying as little as 0.5–1% according to environmental demands suggests that its operations are mainly intrinsic in nature (Raichle, 2006). Maintaining, interpreting, and predicting environmental information and its for demands for behavioural responses may indeed be the most costly of tasks performed by the brain. This could account, for example, for the surprisingly low number of only 10% of synapses in visual cortex carrying incoming information from the outside world: perhaps innate activity is simply more central to the brain's modus operandi. While the exact reasons for the nearly unchanging level of neural activity in the

brain remain unknown, one thing we can surmise from such a pattern of energy consumption is that the brain is never at rest.

5.3 Sparse neural codes

When action potentials are costly, neural signalling must adapt to convey messages in as few spikes as possible without losing information. Energy-efficient sparse coding of neural representations involves the compression of information into a code with less redundancy, employing fewer action potentials and leading to a neural code that is less taxing in a metabolic sense. Experimental evidence of sparse neural codes has been found in a variety of sensory modalities and animal species (see Barth & Poulet, 2012).

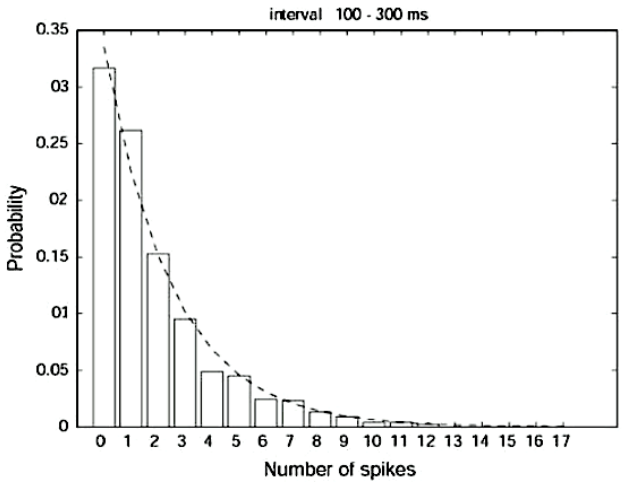

Figure 18. Probability distribution of firing rates in 41 macaque temporal cortical neurons in response to visual stimuli. The dashed line shows the fit to an exponential (L-shaped) distribution. (Franco, Rolls, Aggelopoulos & Jerez, 2007)

Franco et al. (2007) studied the selectiveness of neural activity to visual natural stimuli in macaque inferior temporal cortex. The probability distribution of firing rates formed by

adding the activity of 41 cortical neurons was found to fit an exponential distribution with low firing rates most common. A small number of neurons were found to be highly active, displaying cumulative rates of 5–15 spikes in an interval of 100–300 ms from stimulus onset. For a given mean firing rate, the exponential distribution of firing rates on display here is consistent with both minimising energy consumption and maximising the transmission rate of information.

Sparse coding of inputs has been demonstrated also in the olfactory system. Recording from a set of 177 individually sampled neurons, Poo & Isaacson (2009) report sparse representation of odors in rat primary olfactory cortex. Neural responses to odors were found to be extremely weak, with increases in the average firing rate of active cells of only around 2 Hz and with only some 6% of neurons displaying activity of more than 10 Hz. Low average firing rates together with the scarcity of highly active cells and the low activity fraction of the cell population indicate that odor representations are sparse in olfactory cortex.

Greene, Barrett, Sen & Houghton (2009) compared the receptive fields of neurons in the auditory forebrain of zebra finch to an artificial network constructed under the assumption of sparseness. As songbirds are known to be adept at distinguishing between a large number of different bird songs in their natural environment, recordings of 20 different zebra finch songs were selected as the auditory stimuli to be presented. Recordings of activity in Field L neurons in the zebra finch forebrain show that they are adapted specifically to sparsely encode bird song stimuli, sharing similarities with activity displayed by the network generated specifically to exhibit sparseness.

In addition to sensory regions of the brain employing a sparse coding strategy for incoming stimuli, the degree of sparseness of neural representations has been shown to vary according to contextual factors. Vinje & Gallant (2002) studied the effect of the non-

classical receptive field (nCRF), an area surrounding the classical receptive field (CRF)[16], on information transmission of neurons in primary visual cortex (V1) in macaque. The authors centered visual inputs containing natural scenes on the CRF of V1 neurons, varying stimulus size from one to four times the diameter of CRF. Stimulation of the nCRF was found to significantly increase both the rate of information transmission, the information in bits per spike, and the efficiency of transmission measured as the fraction of available bandwidth used by neurons to transmit information. The authors suggest that the nCRF might tune V1 neurons to match the sparsely distributed informative components in visual natural scenes, resulting in a sparse and highly efficient neural code. Corroborating evidence for this view was later offered by Haider, Krause, Duque, Yu, Touryan, Mazer & McCormick (2010), who replicated these results in both cat primary visual cortex and a simple computational model.

[16] The classical receptive field of a neuron is defined as the spatial region within which a stimulus presented in isolation can evoke a response in the neuron.

6 Discussion

The aim of this study has been to present evidence of the presence and, where possible, extent of biological optimisation in the central nervous system to do with the efficiency of information processing. Especially in the case of wiring minimisation, structural optimisations are seen to reach levels close to their theoretical optima. Oftentimes the solutions shaped by natural selection in response to selective pressures were found to be conspicuously similar to those found in complex technological systems.

The notion of biological optimisation resulting from evolutionary adaptation has met with some scorn in the past (Gould & Lewontin, 1979). Proponents of this approach have stood accused of trying to prove optimality in animals ex post facto and of constructing explanations on the go for the existence and justification of specific traits. While the process of evolutionary adaptation may never produce absolute optimality in a system as complex as the human information processing apparatus, it has been shown to promote biological designs displaying a very high degree of efficiency. The discovery of design rules with general validity in describing the evolution and function of nervous systems speaks loudly to the legitimacy of such an approach and the explanatory power of design principles uncovered in this manner.

Examples of such parsimonious design rules were shown to feature widely in brain anatomy and physiology. The principle of wiring minimisation for one has far-reaching influence on the information processing architecture of the brain. Axon diameter, often accounting for the lion's share of an individual neuron's volume, is optimised to within touching distance of its theoretical minimum set by the buildup of noise in neuronal signalling (Faisal et al., 2005). The comparatively simple nervous system of the nematode worm *C. elegans* is particularly amenable to the study of wiring optimisation, exhibiting

close to optimal placement of neuronal components and hence minimisation of wiring (Pérez-Escudero et al., 2009).

Likewise, the drive for wiring optimisation can be seen to push for the separation of gray and white matter, for a wiring fraction of 3/5 in optimally wired gray matter and for the efficient layout of cortical regions. While the rule of "save wire" is certainly not the only factor promoting these design choices in brain architecture, it retains an almost singular validity in predicting features of neuroanatomy both across a range of species and of organisational levels. Given time and more comprehensive anatomical datasets, wiring optimisation may yet shed light on more specific rules of neural connectivity that could allow for the inference of neuronal function from anatomical data, illuminating the currently ambiguous relationship between brain anatomical and functional networks.

Just as wiring minimisation dictates that neuronal networks be designed so as to conserve biological cabling, so the efficient coding hypothesis first proposed more than five decades ago by Barlow (1961) suggests a similar economical principle at work in neural coding and signalling. With natural scenes containing considerably redundant information, the sensory regions of the nervous system in particular are adapted to take advantage of those redundancies. Human cortex as a whole operates under an extreme paucity of energy, with only a few percent of neurons significantly active at any time.

The sparse coding of neural representations can be seen as a direct corollary of the principle of efficient coding. Such a coding scheme combines representational and metabolic efficiency by recruiting a modest fraction of neurons to encode a single representation, skirting the high energetic costs inherent in fully distributed codes while still achieving a combinatorial capacity far beyond that of metabolically favourable local codes. As would be expected on the basis of these twin efficiencies, sparse neural codes

have been observed in a variety of sensory modalities spanning a range of species (see Barth & Poulet, 2012).

While biological structures and man-made technological systems clearly differ in physical implementation, similarities can be discerned especially at higher levels of organisation. Small-world networks, for example, are found repeatedly at different organisational levels of brain structural and functional networks, just as they are elsewhere in natural, social, and technological domains. Local clusters typical of small-world networks are recombined at multiple levels of brain organisation to produce larger functional units, bringing to mind the modular design of many technological systems from Lego blocks to computer software modules.

Exploitation of the redundancy present in sensory data and the efficient representation of relevant information through the use of energy-efficient sparse neural codes, too, share characteristics with technological solutions employed in the engineering of communications networks. Likewise, component-placement optimisation is a central concern in the design of electronic devices such as digital microchips just as it is in brain anatomy. In both domains an inferior arrangement of components causes excessive wire length and timing issues arising from signalling delays in long-range connections, and wastes energy as a result of unnecessarily long wires.

As for this thesis, the space available has, alas, a prevented a more comprehensive discussion of matters relating to efficient neural information processing. To give but a few examples of topics outside the scope of this study, it is plausible that the unreliable and comparatively slow behaviour of nerve cells has promoted the widespread parallel mode of information processing in the brain. Synaptic plasticity is thought to play an important role in the functioning of both anatomical and functional cortical networks, with significant implications for neural signalling. Likewise, the synchronisation of neuronal activity

across cortical areas has been shown to affect the efficiency of global information transfer in the brain. These themes, alongside numerous others, offer promising avenues for further research into the efficiency of biological information processing.

Contemporary experimental limitations still inhibit enquiry on a number of fronts. Knowledge of wiring minimisation currently rests on a limited number of rather modest sets of anatomical figures. More comprehensive neuroanatomical datasets, perhaps eventually culminating in the holy grail of the human connectome, will in time permit ever more extensive analyses of wiring optimisation at work. In the meantime there is hope for multi-resolution approaches offering a feasible midway point on the road towards the full connectome. Deeper knowledge of human neuroanatomy should also throw more light on the relationship between brain anatomical and functional connectivity. Likewise, the present evidence for stochastic resonance at work in the brain is for the most part indirect, yet it is hoped that more advanced experimental techniques will in due course provide direct evidence of noise intrinsic to the brain playing a constructive role in neural signalling (McDonnell & Abbott, 2009).

The nature of neural coding presents another predicament to contemporary experimental research. For one, the sparsity of neuronal activity in the human cortex complicates the task of finding relevant spiking neurons in the first place. The highly selective firing patterns of neurons in higher-order sensory cortical areas add a further level of difficulty, technological limitations to do with large-scale recordings of the activity of individual nerve cells another. Novel techniques such as optical stimulation and recording of neurons, as well as longitudinal recordings, should in time permit a more detailed view of the nature of sparse neural representations.

Finally, some questions arise for future research. Global efficiency deriving from small-world characteristics of brain networks is associated with individual differences in human

intellectual performance. That link is curiously different for anatomical and functional networks: the average path length correlates positively with intelligence in both cases, yet the number of neuronal connections does so only for anatomical networks. The reasons for this are unclear at present, but this direct link between brain network efficiency and intellectual performance is a distinctly worthy topic of additional study.

Wiring minimisation operates at high efficiency even for optimisation problems of significant computational complexity. What is perhaps most remarkable about this phenomenon is the apparent lack of either a collective or centralised mechanism guiding the process. With the availability of ever larger anatomical datasets in particular, we may soon be able to assess whether or not wiring efficiency eventually reaches a limit at some higher stage of complexity.

Behavioural requirements relating to information processing and neural coding are linked intricately to neuroanatomy by biophysical constraints and provide a rich area of investigation. Limits to the human brain's energy supply confine the brain's "clock speed" to the millisecond range (Attwell & Gibb, 2005). This raises the question of whether faster processing would not confer selective value great enough to warrant its development, or whether perhaps a higher frequency of computation is simply too costly given the sources of energy (food) available during human evolutionary history.

Perhaps the most important question left unanswered is one concerning the possible future application of rules governing efficient brain design. Will knowledge of these seemingly universal principles eventually permit researchers to predict aspects of hitherto uncharted territory in the brain, say, in the domain of higher-order neural representations? We are not yet at that level – and we may never be – but conceivably the exposition of these principles of brain function may in time help to remove that "eternally nagging doubt" (Cherniak,

1994) of whether we really understand how the brain works much better than a bunch of chimpanzees taking apart a radio.

References

Achard, S. & Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Computational Biology, 3*(2), e17.

Aiello, L. C. and Wheeler, P. (1995). The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution. *Current Anthropology, 36*(2), 199-221.

Alexander, R. M. (1996). *Optima for animals.* Princeton, NJ: Princeton University Press.

Alexander, R. M. (2001). Design by numbers. *Nature, 412,* 591.

Allen, C. & Bekoff, M. (1995). Biological function, adaptation, and natural design. *Philosophy of Science, 62*, 609–622.

Alon, U. (2003). Biological networks: The tinkerer as an engineer. *Science, 301,* 1866–1867.

Ames, A. (2000). CNS energy metabolism as related to function. *Brain Research Reviews, 34,* 42–68.

Attwell, D. & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow and Metabolism, 21,* 1133–1145.

Attwell, D., & Gibb, A. (2005). Neuroenergetics and the kinetic design of excitatory synapses. *Nature Reviews Neuroscience, 6*(11), 841-849.

Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory Communication* (pp. 217-234). Cambridge, MA: MIT Press.

Barth, A. L. & Poulet, J. F. A. (2012). Experimental evidence for sparse firing in the neocortex. *Trends in Neurosciences, 35*(6), 345–355.

Bassett, D. S. & Gazzaniga, M. S. (2011). Understanding complexity in the human brain. *Trends in Cognitive Sciences, 15*(5), 200–209.

Bhalla, U. S. (2008). How to record a million synaptic weights in a hippocampal slice. *PLos Computational Biology, 4*(6), e1000098.

Briggs, F. (2009). Organizing principles of cortical layer 6. *Frontiers in Neural Circuits, 4*(3), 1–8.

Brownlee, J. (2007). Satisficing, optimization, and adaptive systems. *CIS Technical Report,* 070305A.

Bullmore, E. & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience, 10*, 186–198.

Bullmore, E. & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience, 13*, 336–349.

Chandler, D. M. & Field, D. J. (2007). Estimates of the information content and dimensionality of natural scenes from proximity distributions. *Journal of the Optical Society of America A, 24*(4), 922–941.

Chen, B. L., Hall, D. H. & Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the Natural Academy of Sciences of the United States, 103*(12), 4723–4728.

Cherniak, C. (1994). Philosophy and computational neuroanatomy. *Philosophical Studies, 73*, 89–107.

Cherniak, C. (1995). Neural component placement. *Trends in Neurosciences, 18,* 522–527.

Cherniak, C. (1996). Component-placement optimization in the brain. *Trends in Neurosciences, 19*(10), 414–415.

Cherniak, C., Mokhtarzada, Z., Rodriguez-Esteban, R. & Changizi, K. (2004). Global optimization of cerebral cortex layout. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 1081–1086.

Cherniak, C. (2012). Neural wiring optimization. In Hofman, M. A. & Falk, D. (eds.), *Evolution of the Primate Brain: From Neuron to Behavior (Progress in Brain Research)* (pp. 361–371). Amsterdam: Elsevier.

Chklovskii, D. B., Schikorski, T. & Stevens, C. F. (2002). Wiring optimization in cortical circuits. *Neuron, 34,* 341–347.

Chklovskii, D. B. & Koulakov, A. A. (2004). Maps in the brain: What can we learn from them? *Annual Review of Neuroscience, 27,* 369–392.

Crespi, B. J. (2000). The evolution of maladaptation. *Heredity, 84*, 623–629.

Erdős, P. & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 5*, 17–61.

Faisal, A. A., White, J. A. & Laughlin, S. B. (2005). Ion-channel noise places limits on the miniaturization of the brain's wiring. *Current Biology, 15,* 1143–1149.

Faisal, A. A. & Laughlin, S. B. (2007). Stochastic simulations on the reliability of action potential propagation in thin axons. *PLoS Computational Biology, 3*(5), e79.

Faisal, A. A., Selen, L. P. J. & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9*, 292–303.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon.

Földiák, P. (2002). Sparse coding in the primate cortex. In Arbib, M. A. (ed.), *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA: MIT Press.

Franco, L., Rolls, E. T., Aggelopoulos, N. C. & Jerez, J. M. (2007). Neuronal selectivity, population sparseness, and ergodicity in the inferior temporal cortex. *Biological Cybernetics, 96*, 547–560.

Futuyma, D. J. (2009). *Evolution*. Sunderland, MA: Sinauer Associates.

Gai, Y., Doiron, B., Kotak, V. & Rinzel, J. (2009). Noise-gated encoding of slow inputs by auditory brain stem neurons with a low-threshold K^+ current. *Journal of Neurophysiology, 102*(6), 3447–3460.

Gammaitoni, L., Hänggi, P., Jung, P. & Marchesoni, P. (2009). Stochastic resonance: A remarkable idea that changed our perception of noise. *The European Physical Journal B, 69*, 1–3.

Gangestad, S. W. & Yeo, R. A. (1997). Behavioral genetic variation, adaptation and maladaptation: An evolutionary perspective. *Trends in Cognitive Sciences, 1*(3), 103–108.

Gilbert, S. F. (2010). *Developmental Biology (9th ed.)*. Sunderland, MA: Sinauer Associates.

Godfrey-Smith, P. (2001). Three kinds of adaptationism. In Orzack, S. H. & Sober, E. (eds.), *Adaptationism and Optimality*. Cambridge: Cambridge University Press.

Godfrey-Smith, P. & Wilkins, J. F. (2006). Adaptationism. In Sarkar, S. & Plutynski, A. (eds.), *A Companion to the Philosophy of Biology (Blackwell Companions to Philosophy)*. Oxford: Wiley-Blackwell.

Golberg, J. L. (2003). How does an axon grow? *Genes & Development, 17*, 941–958.

Gould, S. J. & Lewontin, R. C. (1979). The spandrels of San Marco and the panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London B, 205*, 581–598.

Gould, S. J. & Vrba, E. S. (1982). Exaptation – A missing term in the science of form. *Paleobiology, 8*(1), 4–15.

Graham, D. J., & Field, D. J. (2006). Sparse coding in the neocortex. In J. H. Kaas (Ed.), *Evolution of Nervous Systems*. London: Academic Press.

Greene, G., Barrett, D., Sen, K. & Houghton, C. (2009). Sparse coding of birdsong and receptive field structure in songbirds. *Network, 20*(3), 162–177.

Gross, C. G. (2002). Genealogy of the "grandmother cell". *The Neuroscientist, 8*(5), 512–518.

Haider, B., Krause, M. R., Duque, A., Yu, Y., Touryan, J., Mazer, J. A. & McCormick, D. A. (2010). Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. *Neuron, 65*, 107–121.

Hall, B. K. (1999). *Developmental Biology*. Dordrecht: Kluwer Academic Press.

Hänggi, P. (2002). Stochastic resonance in biology: How noise can enhance detection of weak signals and help improve biological information processing. *ChemPhysChem, 3*, 285–290.

Harris, J. J. & Attwell, D. (2012). The energetics of CNS white matter. *The Journal of Neuroscience, 32*(1), 356–371.

Hartline, D. K. & Colman, D. R. (2007). Rapid conduction and the evolution of giant axons and myelinated fibers. *Current Biology, 17*, R29–R35.

Herculano-Houzel, S., Mota, B., Wong, P. & Kaas, J. H. (2010). Connectivity-driven white matter scaling and folding in primate visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 107*(44), 19008–19013.

Herculano-Houzel, S. (2012). The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proceedings of the National Academy of Sciences of the United States of America, 109* (suppl. 1), 10661–10668.

Hilgetag, C., Burns, G., O'Neill, M., Scannell, J. & Young, M. (2000). Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat. *Philosophical Transactions of the Royal Society of London B, (355)*, 91–110.

Hofman, M. A. (2001). Brain evolution in hominds: Are we at the end of the road? In Falk. D., Gibson, K. (eds.), *Evolutionary Anatomy of the Primate Cerebral Cortex*. Cambridge: Cambridge University Press.

Ikeda, M. & Oka, Y. (2012). The relationship between nerve conduction velocity and fiber morphology during peripheral nerve regeneration. *Brain and Behavior, 2*(4), 382–390.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transaction on Neural Networks, 14*(6), 1569–1572.

Kaiser, M. (2007). Brain architecture: A design for natural computation. *Philosophical Transactions of the Royal Society of London A, 365*, 3033–3045.

Kaiser, M. & Hilgetag, C. C. (2004). Modelling the development of cortical systems networks. *Neurocomputing, 58–60*, 297–302.

Klyachko, V. A. & Stevens, C. F. (2003). Connectivity optimization and the positioning of cortical areas. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 7937–7941.

Kosinski, R. A. & Zaremba, M. (2007). Dynamics of the model of the *Caenorhabditis elegans* neural network. *Acta Physica Polonica B, 38*(6), 2201–2210.

Kourkoulas-Chondrorizos, A. (2012). *Online optimisation of information transmission in stochastic spiking neural systems* (Doctoral dissertation). Retrieved from University of Edinburgh Research Archive (Publication no. EP/E002005/1).

Laughlin, S. B. & Sejnowski, T. J. (2003). Communication in neuronal networks. *Science, 301,* 1870–1874.

Lennie, P. (2003). The cost of cortical computation. *Current Biology, 13,* 493–497.

Levy, W. B. & Baxter, R. A. (1996). Energy efficient neural codes. *Neural Computation, 8,* 531–543.

Leung, M., Williams, P., Benedetto, A., Au, C., Helmcke, K., Aschner, M. & Meyer, J. (2008). *Caenorhabditis elegans*: An emerging model in biomedical and environmental toxicology. *Toxicological Sciences, 106*(1), 5–28.

Lewens, T. (2002). Adaptationism and engineering. *Biology and Philosophy, 17*, 1–31.

Li, Y., Liu, Y., Li, J., Qin, W., Li, K., Yu, C. & Jiang, T. (2009). Brain anatomical network and intelligence. *PLoS Computational Biology, 5*(5), e10000395.

Mandal, S. (2009). *Collective analog bioelectric computation* (Doctoral dissertation). Retrieved from DSpace@MIT (Publication no. 1721.1/52801).

Mark, R. (1996). Architecture and evolution. *American Scientist, 84*(4), 383–389.

Matthen, M. & Ariew, A. (2002). Two ways of thinking about fitness and natural selection. *The Journal of Philosophy, 99*(2), 55–83.

Maynard Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics, 9*, 31–56.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D. & Wolpert, L. (1985). Developmental constraints and evolution: A perspective from the Mountain Lake Conference on Development and Evolution. *Quarterly Review of Biology, 60*(3), 265–287.

Mayr, E. (1983). How to carry out the adaptationist program? *The American Naturalist, 121*(3), 324–334.

Mayr, E. (2004). *What Makes Biology Unique?* Cambridge: Cambridge University Press.

McDonnell, M. D. & Abbott, D. (2009). What Is Stochastic Resonance? Definitions, Misconceptions, Debates, and Its Relevance to Biology. *PLoS Computational Biology, 5*(5), e1000348.

McDonnell, M. D. & Ward, L. M. (2011). The benefits of noise in neural systems: Bridging theory and experiment. *Nature Reviews Neuroscience, 12*, 415–425.

McIntosh, A. R., Kovacevic, N., Lippe, S., Garrett, D., Grady, C. & Jirsa, V. (2010). The development of a noisy brain. *Archives Italiennes de Biologie, 148*, 323–337.

Moss, F., Ward, L. M., Sannita, W. G. (2004). Stochastic resonance and sensory information processing: A tutorial and review of application. *Clinical Neurophysiology, 115*, 267–281.

Neishabouri, A. M. & Faisal, A. A. (2011). The metabolic efficiency of myelinated vs unmyelinated axons. *BMC Neuroscience, 12* (suppl. 1).

Neumann, J. von (1958). *The Computer and the Brain*. New Haven, CT: Yale University Press.

Nielsen, R. (2009). Adaptationism – 30 years after Gould and Lewontin. *Evolution, 63*(10), 2487–2490.

Noctor, S. C., Martínez–Cerdeño, V. & Kriegstein, A. R. (2007). Contribution of intermediate progenitor cells to cortical histogenesis. *Archives of Neurology, 64*(5), 639–642.

Olshausen, B. A. & Field, D. J. (2005). How close are we to understanding V1? *Neural Computation, 17*, 1665–1699.

Orzack, S. H. & Sober, E. (1996). How to formulate and test adaptationism. *The American Naturalist, 148*(1), 202–210.

Parker, G. A. & Maynard Smith, J. (1990). Optimality theory in evolutionary biology. *Nature, 348*, 27–33.

Pérez-Escudero, A. & Polavieja, G. G. de la (2007). Optimally wired subnetwork determines neuroanatomy of *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America, 104*(43), 17180–17185.

Pérez-Escudero, A., Rivera-Alba, M. & Polavieja, G. G. de la (2009). Structure of deviations from optimality in biological systems. *Proceedings of the National Academy of Sciences of the United States of America,106*(48), 20544–20549.

Pigliucci, M. & Kaplan, J. (2000). The fall and rise of Dr Pangloss: adapatationism and the *Spandrels* paper 20 years later. *Trends in Ecology and Evolution, 15*(2), 66–70.

Pigliucci, M. & Kaplan, J. (2006). *Making Sense of Evolution: The Conceptual Foundations of Evolutionary Biology.* Chicago, IL: The University of Chicago Press.

Potochnik, A. (2009). Optimality modeling in a suboptimal world. *Biology and Philosophy*, *24*, 183–197.

Raichle, M. E. (2006). The brain's dark energy. *Science, 314,* 1249.

Raichle, M. E. & Mintun, M. A. (2006). Brain work and brain imaging. *Annual Review of Neuroscience, 29,* 449–476.

Raj, A. & Chen, Y. (2011). The wiring economy principle: Connectivity determines anatomy in the human brain. *PLoS One, 6*(9), e14832.

Ramón y Cajal, S. (1995). *Histology of the Nervous System of Man and Vertebrates*. Trans. Swanson, N. & Swanson, L. New York: Oxford University Press.

Reeve, H. K. & Sherman, P. W. (1993). Adaptation and the goals of evolutionary research. *The Quarterly Review of Biology, 68*(1), 1–32.

Rolls, E. T. & Treves, A. (2011). The neuronal encoding of information in the brain. *Progress in Neurobiology, 95*, 448–490.

Roth, G. & Dicke, U. (2005). Evolution of the brain and intelligence. *Trends in Cognitive Sciences, 9,* 250–257.

Rudolph, M. & Dextexhe, A. (2001). Do neocortical pyramidal neurons display stochastic resonance? *Journal of Computational Neuroscience, 11*, 19–42.

Sarpeshkar, R. (1998). Analog versus digital: Extrapolating from electronics to neurobiology. *Neural Computation*, 10, 1601–1638.

Sarpeshkar, R. & O'Halloran, M. (2002). Scalable hybrid computation with spikes. *Neural Computation, 14*, 2003–2038.

Sarpeshkar, R. (2012). Universal principles for ultra low power and energy efficient design. *IEEE Transactions on circuits and systems, 59*(4), 193–198.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27,* 379–423.

Shoham, S., O'Connor, D. H. & Segev, R. (2006). How silent is the brain: Is there a "dark matter" problem in neuroscience? *Journal of Comparative Physiology A, 192*(8), 777–784.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.

Simoncelli, E. P. & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24*, 1193–1216.

Schlosser, G. (2007). Functional and developmental constraints on life-cycle evolution. In Sansom, R. & Brandon, R. N. (eds.), *Integrating Evolution and Development*. Cambridge, MA: The MIT Press.

Schwenk, K. (1995). A utilitarian approach to evolutionary constraint. *Zoology, 98*, 251–262.

Sporns, O., Chialvo, D. R., Kaiser, M. & Hilgetag, K. C. (2004). Organization, development and function of complex brain networks. Trends in Cognitive Sciences, 8, 418–425.

Sporns, O., Tononi, G. & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology, 1*(4), e42.

Sporns, O., & Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics, 2,* 145–162.

Stephan, K., Hilgetag, C., Burns, G., O'Neill, M., Young, M. & Kötter, R. (2000). Computational analysis of functional connectivity between areas of primate cerebral cortex. *Philosophical Transactions of the Royal Society of London B, 355*, 111–126.

Stewart, C. (1993). The powers and pitfalls of parsimony. *Nature, 361*, 603–607.

Sutherland, W. J. (2005). The best solution. *Nature, 435,* 569.

Traynelis, S. F. & Jaramillo, F. (1998). Getting the most out of noise in the central nervous system. *Trends in Neurosciences, 21,* 137–145.

Heuvel, M. P. van den, Stam, C. J., Kahn, R. S. & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *The Journal of Neuroscience, 29*(23), 7619–7624.

Vilarroya, O. (2002). "Two" many optimalities. *Biology and Philosophy, 17*, 251–270.

Vinje, W. E. & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *The Journal of Neuroscience, 22*(7), 2904–2915.

Wang, S., Shultz, J., Burish, M., Harrison, K., Hof, P., Towns, L., Wagers, M. & Wyatt, K. (2008). Functional trade-offs in white matter axonal scaling. *The Journal of Neuroscience, 28*(15), 1–10.

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393,* 440– 442.

Wen, Q. & Chklovskii, D. B. (2005). Segregation of the brain into gray and white matter: A design minimizing conduction delays. *PLoS Computational Biology, 1,* 0617–0630.

White, J. A., Rubinstein, J. T., & Kay, A. R. (2000). Channel noise in neurons. *Trends in Neurosciences, 23*(3), 131-137.

Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Willmore, B. & Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Network: Computation in Neural Systems, 12*(3), 255–270.

Wood, H. (2011). A rapid e-volution. *Nature Reviews Neurology, 7*, 415.

Wright, S. (1932). The roles mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the VI International Congress of Genetics, 1*, 356–366.

Young, M. P. & Scannell, J. W. (1996). Component-placement optimization in the brain. *Trends in Neurosciences, 19*(10), 413–414.

Zhang, K. & Sejnowski, T. J. (2000). A universal scaling law between gray matter and white matter of cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America, 97,* 5621–5626.