SIR DAVID JOHN CAMERON MACKAY FRS
22 April 1967 — 14 April 2016
SIR DAVID JOHN CAMERON MACKAY FRS

22 April 1967 — 14 April 2016

Elected FRS 2009

BY MALCOLM LONGAIR, CBE, FRS, FRSE1,* AND
MICHAEL CATES, FRS, FRSE2

1 Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, UK
2 Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

David MacKay was a true polymath who made pioneering contributions to information theory, inference and learning algorithms. He was a founder of the modern approach to information theory, combining Bayesian inference with artificial neural network algorithms to allow rational decision making by computer. His major achievements include reliable computation with unreliable hardware, in particular in approaching the Shannon limit using enhancements of Gallager codes. He developed communication systems for the disabled, including the Dasher code which he made freely available. He was the author of the influential book Sustainable Energy: Without the Hot Air. In 2009 he was appointed Chief Scientific Advisor to the Department of Energy and Climate Change. From 2003 he was Professor of Natural Philosophy in the Cavendish Laboratory, Cambridge, and in 2013 he was appointed to the first Regius Professorship of Engineering in the Engineering Department of Cambridge University. He was appointed Knight Bachelor in the 2016 New Year’s Honours List ‘for services to Scientific Advice in Government and Science Outreach’, but lost his battle with stomach cancer soon afterwards at the age of 48.

FAMILY BACKGROUND AND EDUCATION

David MacKay was born on 22 April 1967, the fifth and youngest child of Donald MacCrimmon MacKay and Valerie MacKay (née Wood). Donald MacKay was a son of the

*msl1000@cam.ac.uk

© 2017 The Author(s)

http://dx.doi.org/10.1098/rsbm.2017.0013 445 Published by the Royal Society
manse—he was born in the Free Church manse at the northern fishing village of Lybster in Caithness in 1922. After Wick High School, Donald graduated from St Andrews University in 1943 with a degree in physics. He was then recruited into the Admiralty for three years, where he worked on the development of radar. In 1946, he moved to King’s College London (KCL) where he completed his PhD in 1951 on ‘The application of electronic principles to the solution of differential equations in physics’ (MacKay, 1951). He was a lecturer in physics at KCL until 1960 when he and his growing family moved to the University of Keele in Staffordshire, where he helped establish the renowned Department of Communication and Neuroscience. This was a highly successful, purely research department sponsored by Granada television. MacKay senior interpreted the remit of the Department to include how communication works in the brain, in particular with the eyes and ears. His research interests led to his book Information, mechanism and meaning (MacKay, 1969). Donald was a committed Christian and delivered the 1986 Gifford Lectures on Natural Theology at the University of Glasgow with the title Under our own microscope: what brain science has to say about human nature. He died the following year from lymphoma. His lectures, edited by his wife Valerie, were posthumously published under the title Behind the eye (MacKay, 1991).

Valerie Wood, born in 1934 in London, was one of Donald’s students at KCL. They married in 1955 and Valerie soon had her hands full with their five children. David was the youngest of these; once he was two years old, Valerie began part-time secondary school teaching. She later carried out research for a PhD on the McCollough effect, a visual illusion in which, after exposure to red stripes in one direction and green in another, black and white stripes appear to be green or red according to their orientation. Not long after she completed her PhD, Donald’s cancer was diagnosed and she spent most of the following period looking after him and then her own father.

In order of birth, David’s siblings were: Robert, now Professor of Mathematics at the University of Warwick (FRS 2000); Eleanor, a speech therapist who worked in Lincoln; Janet, a General Practitioner and Senior Lecturer in Medical Education at Keele University; and Margaret, business manager for the Latimer Trust, a Christian publishing house. Evidently, David came from an intellectually lively and stimulating background, his father’s research interests surely making a significant impact upon David’s future professional interests in information theory and communication. A profound love of music was also instilled at an early age; his tastes were eclectic, embracing excellence in every genre from the romantic period, through modernist minimalism, to rock music.

At primary school, David played soccer and rounders. For his secondary school education, he attended the local grammar school at Newcastle-under-Lyme in Staffordshire, where he took 10 O-levels, six A-levels and two S-levels, all with the highest grades. He played judo, hockey, violin and viola, and did a lot of cycling. In 1985, when he was 18, he represented Britain at the International Physics Olympiad in Yugoslavia, where he won the silver medal overall and the first prize for experimental work.

In 1986, he won an Entrance Scholarship, and subsequently a Senior Scholarship, to study at Trinity College, Cambridge, where he pursued the course in natural sciences with characteristic enthusiasm and success. He enjoyed his time at Trinity, living in the Great Court for a year and also looking after a punt and playing croquet. In the vacations, part of his time was spent climbing mountains in Wales and Scotland, a recreation he continued to pursue throughout adulthood. After his final physics examinations, he rowed for a couple of weeks, went down four in the fifth First and Third Trinity boat and experienced the May week ‘boatie’
dinner at Trinity. David was not someone to miss out on what undergraduate life in Cambridge had to offer!

His first research work took place during the 1986 Long Vacation at the Royal Signals and Radar Establishment (RSRE), Malvern, where he was given the task of testing high-precision digitizers statistically. These were placed in uniformly distributed random voltages and then the distribution of digital read-outs was observed.¹ That summer he also sent a contribution to an Institute of Physics magazine giving a solution to the problem of constructing a spiral mirror. Other topics that amused him at that time were the construction of astrolabes and making a learning-and-prediction programme which attempted to predict the next digit in a human-generated binary sequence.

In 1987 he attended the first International Conference on Neural Networks (ICNN) in San Diego. He then returned to RSRE Malvern in the Long Vacation, this time under the supervision of John Bridle. David worked on a modification of ‘Netspeak’, described as ‘a neural network that learns to read aloud’. This involved feeding the old outputs back as extra inputs to the hidden layer, the idea being that this extra contextual information should improve performance, as indeed was found to be the case.² This work was presented as a poster at a meeting at RSRE in late 1987. David also exhibited a poster on the idea of exploring constant energy surfaces in weight space for neural networks with more parameters than data points. He always acknowledged how important this early work at Malvern was in stimulating him to pursue his research career in physics and information science.

David graduated with a BA degree in Natural Sciences (Physics and Theoretical Physics) from Trinity College in 1988, having obtained first class honours in each year of the Tripos examinations. At some point during his undergraduate days, he came to reject the deep-rooted Christian beliefs with which he had been brought up. However, he never lost the work ethic, positive attitude and commitment to the betterment of mankind that went with them. He found these equally compatible with an atheist worldview that he maintained for the rest of his life.

PhD studies at Caltech

Following his BA degree, David won a Fulbright Scholarship to study for a PhD in Computation and Neural Systems at the California Institute of Technology (Caltech) in Pasadena under the supervision of John Hopfield. His first research project was to investigate Linsker’s model of the spontaneous emergence of tuned neuron response in the visual cortex. After a couple of months he teamed up with Kenneth Miller and they published a series of papers on this topic.³ David also worked with Seymour Benzer on the expression of beta-galactosidase in genetically modified Drosophila. He studied the polka-dot expression patterns in the brains of larvae and adults and found some interesting patterns, including one strain in which the stained cells appeared to be associated with the optic chiasm of the adult. A member of Benzer’s group showed him a paper by Delbruck on estimating the mutation rate of an exponentially growing bacterial population. The estimator was clearly unreliable, but at the time David still had not fully absorbed the Bayesian concepts that would later allow him to find unerringly optimal answers to even the most complex statistical questions. Instead, he suggested a collection of ad-hoc estimators, some of which had improved sampling properties.

¹ Numbers in this form refer to the bibliography at the end of the text.
Soon after, David began to use Bayesian and Maximum Entropy methods for Hopfield networks and solved a special case, a cycle-free network. By this time, he knew that Bayesian methods could help with complexity control and model comparison. In 1990, he demonstrated that minimum description length (MDL) and Bayesian methods were equivalent. His brother Robert recalls David’s ‘evangelistic fervour for Bayesian inference’ on a visit to Caltech in November–December 1990. He also wrote software to carry out mixture-of-Gaussians modelling and density modelling with other distributions. For each model, he fitted a Gaussian approximation to the posterior and computed the evidence.

Attending the conference Maxent 90 was a key moment. There, David consolidated what he knew about Bayesian complexity control, and David Robinson explained many of the details involved in hierarchical modelling for image reconstruction. At the end of 1990, David (MacKay) began implementing Bayesian methods for neural networks. By April 1991 he had written two seminal papers on regression problems. ‘Bayesian interpolation’ (7) and ‘A practical Bayesian framework for backpropagation networks’ (8) were presented at the Snowbird conference in Utah, his first exposure at a major meeting. Radford Neal got in touch with David in April 1991 and this began his long-lasting discussion of Bayesian methods for neural nets, and many other topics.

John Hopfield has described as follows the positive impact David made at Caltech:

No description of David at Caltech would be complete without mentioning the effect of his clarity of thought and exposition on those around him. He was a concise and constructive critic of the research of others, a joyful participant in wide-ranging scholarly debate. (personal communication)

David completed his PhD ‘Bayesian methods for adaptive models’ in three years and four months—exceptionally rapid for a Caltech PhD (6). This became a standard resource for those new to this subject, and remained so at least until David’s own comprehensive book on information theory was published (23).

In California, David discovered that he really cared about green politics and became the founding chair of the Caltech Environmental Task Force. He bought a car, essential in California, but decided that he did not want to live in a car-dominated society. Although Caltech was outstanding academically, he resolved to return to the green spaces of Cambridge.

Cambridge and Darwin College

The chance to do so came in 1992 when David was awarded a Royal Society Smithson Research Fellowship, which he held at Darwin College, one of the graduate student colleges founded in Cambridge in the 1960s. His first academic base was the Radio Astronomy Group in the Cavendish Laboratory where Stephen Gull and his colleagues were pioneering and exploiting Maximum Entropy and Bayesian techniques in many different fields, including what would now be called (astronomical) ‘big data’. On his return, he bought a house close to Girton College and a sufficiency of bicycles, including a folding Brompton and a tandem (figure 1). The house became the testing ground for a number of environmental sustainability experiments, later to be recorded in his influential book Sustainable energy: without the hot air (31).
In Cambridge, David attracted further attention for his PhD work by writing review papers, including one on optimization versus ‘integrating out’ hyperparameters (12). His interests spanned an enormous range of topics. After discussing with Peter Brown how IBM’s ‘smoothing’ method worked in language modelling, David worked out a Bayesian hierarchical model alternative. He also started working on latent variable models for proteins, and studied the immense chromatic aberration in the human eye. Intriguingly, this topic was related to his mother’s PhD research on the McCollough effect (MacKay, 1978) and to his parents’ joint research papers, for example MacKay and MacKay (1977). His description of the remarkable ability of the eye to overcome these aberrations, and simple experiments to demonstrate the effects, are contained in sect. 46.3 of his book *Information theory, inference and learning algorithms* (23).

In 1993, his paper using an Automatic Relevance Determination (ARD) model won the American Society of Heating, Refrigeration and Air-Conditioning Engineers prediction competition (9). The competition attracted 150 entrants, who attempted to predict the unseen power loads in buildings from weather and solar radiation data. David’s ARD prior helped select the relevant variables from the large number of possible inputs, and his Bayesian neural network algorithm won the competition by some margin. He was justifiably proud of this achievement, which showed how substantial energy footprint reduction could be gained by making the most of readily collected information, anticipating some of the ‘smart cities’ initiatives that have taken place since then.

In the same year, David was asked to investigate a cryptanalysis problem which introduced him to information and coding theory. He came up with a variational free energy minimization algorithm, motivated by Cheeseman’s description of solving the colouring problem presented
at Maxent 90 (11). This marked the start of his work on ‘ensemble adapting’ methods, also known as variational methods.

David discussed with Radford Neal how to use this approach to solve new problems, one challenge being to get closer to the Shannon limit for error-correcting codes. They initially worked from the idea that it might help to introduce redundancy by using a sparse source, rather than the conventional approach of transmitting non-redundant source bits unchanged, and then adding redundancy in the form of additional parity check bits. David implemented and tested these MN (MacKay–Neal) codes, finding encouraging but not amazing performance, though later he did see these as leading to a more general class that also includes ‘repeat-accumulate codes’.

Exploring variations on MN codes, David devised a class of codes with sparse parity check matrices, and a way of decoding messages transmitted using such codes that iteratively updated probabilities, in a way that tried to avoid double counting of information.

Neal recognized that this decoding method was the same as Pearl’s belief-propagation method for his ‘Bayesian networks’, also known as ‘directed graphical models’, except applied to a problem where it gives an approximate rather than exact solution, because the existence of ‘loops’ prevents complete avoidance of double counting. This method is now generally known as loopy belief propagation, and has been used for approximate inference in problems other than coding. In fact, Gallager had devised the same method, and used it for decoding this same type of Low Density Parity Check (LDPC) code, many years before Pearl (Gallager, 1962).

Data are encoded with the LDPC code as blocks of bits (0s and 1s), with redundancy introduced by requiring that transmitted blocks must satisfy a set of parity checks on various subsets of bits. Random codes with dense parity checks, on subsets of about half the bits, would work very well, if it weren’t that decoding them is computationally impractical. Many classes of codes use parity checks with special structure for which the decoding computation is feasible. LDPC codes are usually constructed randomly, but using a scheme that leads to the matrix of parity checks being sparse, which is crucial for decoding by loopy belief propagation to work well, despite it not being exact.

The outcome of the joint work with Neal was a major paper, *Near Shannon limit performance of low density parity check codes* (13). This paper showed that LDPC codes could compete with the then state-of-the-art, known as Turbo codes, in the race to approach channel capacity. It also introduced irregular LDPC codes, with some columns of the parity check matrix having three 1s and others having only two 1s. This turned out to be a crucial innovation, though not entirely understood at the time.

David’s work with Neal led to a revival of interest in Gallager’s codes, and in 1999 Gallager was awarded the Technion Harvey Prize in Science and Technology. With Matthew Davey, David introduced enhancements to Gallager’s codes that could outperform the Turbo codes with record-breaking performance (17). Overall, this body of research represented a major breakthrough in understanding how information could be accurately transmitted through noisy communication links, approaching close to the fundamental limits imposed by Shannon’s Theorem. These codes are now used widely in devices such as computer disk drives, mobile phone networks, digital broadcasting and Wi-Fi.

In 1996, David heard Terrence (Terry) Sejnowski describe Independent Component Analysis (ICA) (Makeig et al., 1996) and shortly afterwards he worked out a much simpler maximum likelihood derivation of ICA (14). In Erice in 1996, he had the idea of combining
Jordan and Jaakkola’s variational methods with Gaussian process classifiers and this was implemented by his student Mark Gibbs (15, 20). In 1997, he wrote a paper on ensemble learning for hidden Markov models (16).

Throughout these years, David played a full role in the life of Darwin College. On his appointment as a Fellow in 1992, he became a member of the College Council, a role which he continued to play until 2008. He served on the Computer, Family, Library and Music Committees, and from 2000 to 2003 was Librarian of the College. In 2002, he co-organized the Annual Darwin College Lecture series with Alan Blackwell on the topic of ‘Power’ (29). This marked his first effort to create an integrated discourse across all aspects of sustainable energy, including the political realities of making a significant impact on the global energy crisis. The lectures, which were by eminent speakers and addressed topics even more broad-ranging than the energy agenda, had a profound and lasting influence on many in the audience.

In 1993, David started regularly playing Ultimate Frisbee on Jesus Green with what he described as ‘a nice bunch of psychologists’. This team sport was very much to his taste. As described on the Cambridge Ultimate website,

Unique to Ultimate, and central to individual and team conduct, is the underlying Spirit of the Game, which embodies the sportsmanship which has sadly been lost from other sports. Players undertake to be competitive but fair and truthful, physical but careful, intense but friendly and courteous.

From 1994 to 2008, David was the Senior Treasurer of Cambridge Ultimate.

FROM UNIVERSITY LECTURER TO PROFESSOR OF NATURAL PHILOSOPHY

David’s appointment to a university lectureship in the Cavendish Laboratory in 1995 was widely applauded and he soon set up his own Inference Group, which was to attract a number of brilliant students. They went on to apply Bayesian probability theory and other approaches in a very wide range of different fields. His meteoric rise through the academic hierarchy was marked by his promotion to a readership in 1999 and to a professorship in 2003, for both of which he chose the title ‘in Natural Philosophy’. He was an outstanding lecturer, his fourth year (Part III) courses on ‘Information theory, pattern recognition and neural networks’ and ‘Materials, electronics and renewable energy’ proving to be particularly popular with both students and staff.

Information theory, inference and learning algorithms

Between 1995 and 2003 David wrote his pioneering 640-page textbook on Information theory, inference and learning algorithms (23). This had begun as a tiny, elegant book of eight chapters corresponding to the eight lectures of his course on information theory. The book gradually expanded as the course became a full 16-lecture course in the Part III year of the physics syllabus, the additional chapters including material on neural networks, variational methods and Monte Carlo methods. As he acknowledged, he borrowed heavily from the expositions of David Spiegelhalter, Walter (Wally) Gilks and Radford Neal. The book’s growth was partly driven by complaints that the original brief expositions were too brief and so he felt obliged to fill in the omitted steps and arguments. He then received complaints about the brevity of the filled-in steps and arguments, so these too had to be expanded. The final result
Biographical Memoirs

was a piece of sustained pedagogy, covering sophisticated technical material, that remains unrivalled. David also commented that

The book also grew because of my lack of self-control—I recklessly added new topics to the book. For me, everything is connected, and it was great fun to include all the things I was interested in, for example, my paper on evolution, sex, and information theory. Rather than go through the inevitable tribulations of submitting it to a conventional journal, just slip it in the book!

A third reason the book grew was that it was written at a time of great change in information theory. When David started the book, the state-of-the-art in error-correction involved use of the Jet Propulsion Laboratory’s (JPL’s) hugely expensive Galileo code. By the time he had finished the book, the codes he wrote with Matthew Davey had matched the performance of Galileo, as had Repeat-Accumulate and Turbo codes. Another revolution at the same time was the invention of digital fountain codes (26), which were squeezed in as a final chapter just as the book was about to be sent to Cambridge University Press for publication in 2003. All these new developments fitted perfectly into the book, since its overarching theme was the connections between machine learning and information theory. The book is a classic. Its title echoes that of his father’s book (MacKay, 1969). As in all his writings, the book was freely available to download from his website.

Following a casual conversation during lunch at Darwin College, David developed a major collaboration with Harry Bhadeshia (FRS 1998) of the Materials Science and Metallurgy Department in Cambridge, whose area of expertise is physical metallurgy. Bhadeshia was interested in understanding and applying neural network methods to model the relationship between the chemical composition, structure and complex properties in steels. These properties, such as toughness and fatigue, can be measured routinely and used in the safe design of engineering structures. They are, however, so complex that there is no adequate theory that can predict them, except in all but the simplest of cases that are not technologically relevant. Given their vastly different backgrounds, it took time and patience to establish a common language, but the collaboration turned out to be productive, leading to a series of 14 joint papers, whose significance is described by Bhadeshia in his review paper of 1999 (Bhadeshia, 1999). David helped in particular to implement his Bayesian and evidence frameworks, which were seminal in avoiding the overfitting problem, leading to some remarkable quantitative predictions even when faced with myriads of input variables.

He introduced the concept of modelling uncertainty, which is independent of the noise that is experienced when an experiment is repeated. This made the models safe to apply because it helped identify domains within the input space where knowledge is sparse or where estimates are made far beyond the known parameters. When combined with thermodynamics and kinetic theory, the work enabled, and continues to enable, the creation of novel alloys to be dramatically accelerated through a rational, evidence-based process that made maximum use of information and thereby reduced the resources needed for the task. Examples include fire-resistant steel, alloys for futuristic power plants and alloys used in the cast condition. These successes led to David’s ‘bigback’ algorithm being adopted in materials research in many parts of the world; more generally, the strategy of combining mechanistic or thermodynamic modelling with informatic, ‘data-mining’ approaches has been very widely copied since then as a way to speed up materials discovery.
The Sally Clark case

David’s expertise in statistics, combined with his passion for fairness and probity, resulted in his involvement in the Sally Clark case. She was a solicitor who, with her solicitor husband, Stephen, had three sons. She was convicted of murdering her first two babies, a few weeks after their births, despite maintaining throughout that the babies had tragically died for no known reason. She was sent to prison to serve two life sentences for murder after an influential ‘expert’ told the jury that it was too improbable statistically that two children in one family could possibly die of natural causes so soon after birth. David was horrified by the blatant misuse of statistics, which had formed a central part of the prosecution’s case. In addition, David was motivated to work on this case as he had close family friends who had similarly lost two infants to Sudden Infant Death Syndrome (SIDS) and personally witnessed the devastation this causes. David demolished the incorrect statistical arguments and then became involved in the campaign for Sally Clark’s release. After more than three years in prison, and five years of fighting within the legal system, Sally was cleared by the Court of Appeal in January 2003.

In July 2005, Sir Roy Meadow, who had testified that the chance of two children from an affluent family suffering SIDS was 1 in 73 million, was found guilty of serious professional misconduct and struck off the register by the General Medical Council (GMC). His calculation deliberately excluded the possibility that SIDS could have a medical cause that would lead to correlations in its occurrence among siblings; rather, he considered SIDS to be a label for rare, non-medical causes of deaths of which murder was among the least unlikely. This GMC decision was later overturned by the High Court. Sadly, Sally Clark died in March 2007, having never fully recovered from the trauma of this dreadful miscarriage of justice.

Dasher

In 1997, while David and Michael Lewicki were on a bus at the Neural Information Processing Systems (NIPS) conference, they reflected on the poor performance of text entry on handheld computers. They were aware that better language models could make it easier for users to type the most likely letters, the now-familiar principle of predictive text. They knew that arithmetic coding models could provide a straightforward prediction mechanism, and that small keyboards might become more efficient if the most likely keys were enlarged in proportion to arithmetic coding ratios so that they were easier to press.

Lewicki was sceptical about the usability of a keyboard with keys jumping around as they were resized, but David continued to explore ideas for illustrating arithmetic coding, as described in his Information theory textbook.2 He constructed a series of interactive software demonstrations of arithmetic coding, culminating in the ‘Dasher’ program, where the user could type by choosing the next most likely characters as they appeared at the right hand side of the screen, or even continuously steer through an abstract landscape of possible texts. The visual impact is so distinctive that a written description cannot do it justice—interested readers should seek one of the many video demonstrations or, even better, use Dasher themselves.

David’s calculations suggested that Dasher would be so fast that it could replace altogether the traditional QWERTY keyboard, which then had no predictive capability. His graduate student David Ward re-engineered his original simple demonstration of bigram frequencies into a robust product with a sophisticated adaptive model. Lunchtime conversations with

2 See figs. 6.1 to 6.5, pages 112 ff, adapted from Witten et al. (Witten et al., 1987).
his Darwin College colleague Alan Blackwell convinced the two of them that empirical evidence would be needed before a *Nature* publication could announce that the keyboard was obsolete. Sadly, experiments with volunteers soon found that the theoretical performance was hard to achieve in practice—the two Davids were by far the fastest Dasher users. Important refinements were introduced, including the ability to go backwards after a mistake as well as a modification so that the character labels on the screen would arrange themselves into readable words (19).

Having learnt that human performance, rather than statistical modelling, was the main constraint on Dasher, the focus of the project moved towards assisting those who were unable to use conventional keyboards. A version of Dasher was created that allowed rapid text entry through eye movements alone—finally justifying the deferred *Nature* publication (22). In 2003 David’s team demonstrated a version controlled simply by breathing, but still able to write at 12 words per minute (28). With Chris Ball, David developed several versions of ‘Button Dasher’ requiring only yes/no decisions (24). In 2005 David demonstrated that he could write at 10 words per minute with one button. Examples of the various ways in which Dasher can be used to input text are illustrated in figure 2.

In 2005, at the ‘Closing the Gap’ meeting, David invented a way in which a totally blind person, having control of only one or two buttons, could plausibly communicate using a Dasher-like system. This opened up remarkable possibilities for severely debilitated patients. Language models for Dasher have now been constructed for more than 100 languages, and versions of the software created for many different computers and operating systems. As with all his work, he made the source code open to all who wanted to use or modify it, and it continues to be maintained by volunteers around the world, including professional software engineers who themselves use Dasher to overcome physical disabilities. Dasher was
made available for the Android market and the iPhone, iPod touch and iPad in 2010. The programme continues to be widely available at no cost, and the symposium in Cambridge devoted to David's work shortly before he died included moving tributes from Dasher users around the world, composed using Dasher itself.\(^3\)

**New directions**

In 2002 the Gatsby charitable foundation awarded David a Senior Research Fellowship to allow him to devote more time to research. As should already be apparent, he had an insatiable curiosity for distinctive and unusual problems. The following list of his interests from the early 2000s is incomplete, but may provide some insight into the extensive nature of his imagination.

- In December 2000 his inference research group won Hopfield and Brody’s ‘mouse brain’ competition. There were two competitions. For competition A, the challenge was to explain how a simulated ‘mouse brain’ made up of about a thousand neurons performed speech recognition; competition B required entrants to construct their own simulated brain, capable of speech recognition on a 10-word vocabulary, like Hopfield and Brody’s. David and Seb Wills won both contests (18).
- In 2001 he started working with Graeme Mitchison on transferring successful ideas from the field of classical error-correcting codes to the neighbouring field of quantum error-correction (25).
- Following the award of his Gatsby Fellowship, he developed new interests in computation using spike-timing networks and Go-playing algorithms.
- In 2003 he developed with his brother, Robert, a mechanistic theory of how biological systems such as actin/myosin convert chemical energy efficiently into work (30). They proposed that molecular motors such as myosin turn chemical energy into mechanical work by a single molecule engine in which a phosphate ion expands a binding pocket, blocked by ADP, with associated lever arm movement. The pocket is permeable to water and so an alternative description is that osmosis to dilute the phosphate ion causes the pocket to swell with water. This proposal has yet to be tested by physiologists, but numerical simulations by Helmut Grubmuller on the related ATPase rotary motor support it.
- At the same time, he became involved in the African Institute for Mathematical Sciences (AIMS), a pan-African network of centres of excellence. The new institute in South Africa provides a one-year course for African graduates in mathematical sciences. David spent roughly eight weeks per year there during the academic years 2003–04, 2004–05 and 2005–06.
- In 2004 he developed an idea jointly with Seb Wills about how the brain works, which he called ‘distributed phase codes’ (Wills, 2004).
- In 2005 Alan Blackwell and he started a joint project called talks.cam, aimed at reducing the labour of being a seminar-series organizer and enhancing cross-disciplinary interactions in the University. They were joined by Phil Cowans, Duncan Simpson and Tom Counsell, and made a system which is now used by all research groups in the University.

\(^3\) See [https://www.youtube.com/watch?v=QxFEUk3J89Q](https://www.youtube.com/watch?v=QxFEUk3J89Q).
• In 2005 Tadashi Tokieda introduced him to a visual illusion called ‘anti-Glass’ or the Trefethen effect (27).

David’s dedication to promoting the value of precise error estimation was invaluable to his many colleagues in Cambridge and beyond. As a personal example, one of us (M.S.L.) was working on obtaining convincing errors on our extensive data-sets concerning the evolution of the most luminous radio galaxies in the Universe at large redshifts. David made available the Bayesian algorithm packages, which he encouraged everyone to use, with the reassurance that they were no more difficult to implement than Gaussian statistics. This indeed proved to be the case and he became a valued co-author of the resulting paper (21).

**SUSTAINABLE ENERGY — WITHOUT THE HOT AIR**

David had a deep interest in the teaching of physics, but held reservations about the traditional mode of teaching through the combination of lecture courses, supervisions and examples classes. He felt that the students were bombarded by difficult and complex ideas and techniques at the expense of being able to make simple order-of-magnitude estimates. Through the Gatsby Foundation, he employed Sanjoy Mahajan as a post-doctoral research associate in the teaching of undergraduate physics. Like David, Mahajan was dedicated to promoting exactly the type of order-of-magnitude reasoning which provides physical insight, but which does not attract much attention in most physics courses. In one of his more striking tests with a set of volunteer undergraduates, Mahajan demonstrated that their answers to physical insight multiple-choice questions were little better than random. The Cavendish Laboratory entrusted him with giving a course in thermodynamics in which he devoted part of the course to carrying out simple order-of-magnitude calculations interactively with the class. Despite his very best intentions and much effort, this experiment did not go down well with either the students or the staff who were too well attuned to the traditional mode of physics teaching as a route to success in the Tripos examinations.

David was, however, soon to find a forum in which the same strategies could be deployed to brilliant effect. This was in quantifying the magnitude of the global energy crisis: indeed David’s book on this topic had as its first working title *You figure it out!* before evolving into the volume finally published in 2008, *Sustainable energy: without the hot air* (31). This remarkable book includes many simple calculations that convincingly demonstrate in quantitative terms both the magnitude of the sustainable energy problem facing all nations, and what different strategies and sources of renewable and other ‘clean’ energies, including (for these purposes) nuclear energy, can contribute to solving it. An example of the simple but compelling type of analysis in the book is shown in figure 3, the left stack showing the energy requirement and the right stack the maximum energy which could be derived from renewable resources. The analysis shows the magnitude of the challenge facing the UK; readers are encouraged to try alternative strategies to make the stacks match, and those from other countries are given the tools to perform the same calculations with local data. The book has been highly praised, not only for its hugely informative, thought-provoking and enabling content, but also for its calm, non-polemical tone. At one point, David, typically, writes:

> Please don’t get me wrong: I’m not trying to be pro-nuclear. I’m just pro-arithmetic.
<table>
<thead>
<tr>
<th>Category</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Defence&quot;</td>
<td>4</td>
</tr>
<tr>
<td>Transporting stuff</td>
<td>12</td>
</tr>
<tr>
<td>Stuff</td>
<td>48</td>
</tr>
<tr>
<td>Food, farming, fertilizer</td>
<td>15</td>
</tr>
<tr>
<td>Gadgets</td>
<td>5</td>
</tr>
<tr>
<td>Light</td>
<td>4</td>
</tr>
<tr>
<td>Heating, cooling</td>
<td>37</td>
</tr>
<tr>
<td>Jet flights</td>
<td>30</td>
</tr>
<tr>
<td>Car</td>
<td>40</td>
</tr>
<tr>
<td>Geothermal</td>
<td>1</td>
</tr>
<tr>
<td>Tide</td>
<td>11</td>
</tr>
<tr>
<td>Wave</td>
<td>4</td>
</tr>
<tr>
<td>Deep offshore wind</td>
<td></td>
</tr>
<tr>
<td>Shallow offshore wind</td>
<td></td>
</tr>
<tr>
<td>Hydro</td>
<td>1.5</td>
</tr>
<tr>
<td>Biomass</td>
<td></td>
</tr>
<tr>
<td>Photovoltaic (PV) farm</td>
<td>50</td>
</tr>
<tr>
<td>PV, 10 m²/p</td>
<td>5</td>
</tr>
<tr>
<td>Solar heating</td>
<td>13</td>
</tr>
<tr>
<td>Wind</td>
<td>20</td>
</tr>
<tr>
<td>Hydro</td>
<td>1.5</td>
</tr>
<tr>
<td>Biomass</td>
<td></td>
</tr>
<tr>
<td>Photovoltaic (PV) farm</td>
<td>50</td>
</tr>
<tr>
<td>PV, 10 m²/p</td>
<td>5</td>
</tr>
<tr>
<td>Solar heating</td>
<td>13</td>
</tr>
<tr>
<td>Wind</td>
<td>20</td>
</tr>
</tbody>
</table>

Figure 3. (Caption overleaf.)
As a result, he was equally respected by organizations such as Greenpeace and Friends of the Earth as by energy companies such as Shell and EDF Energy. The book is freely available online and by 2012 had been translated into French, Japanese, German, Hungarian, Slovak and Polish.

As we write this Memoir, it remains very unclear to what extent humanity will take David's advice about how to 'kick the habit' of fossil fuels, not all of which will be easy to follow. David believed, as do most scientists today, that failure to do so risks wholly negative consequences for the global environment and its human occupants. However, he made a conscious choice not to make these the subject of his book. This was in part because he felt an overwhelming scientific case for man-made climate change had already been made by others, and in part because—in contrast to the issue of energy budgets—the complexities of climate science do not lend themselves to the You figure it out! approach. The resulting pragmatic focus on how rather than why to pursue a sustainable energy policy may have greatly increased the influence of the book.

**Chief Scientific Advisor to DECC**

In 2009 David was elected to the Fellowship of the Royal Society. A consequence of the success and impact of his book on *Sustainable energy* was that he became well known in policy-making circles. In that same year he applied for and was appointed as Chief Scientific Advisor to the Department of Energy and Climate Change (DECC), a post he was to hold for the next five years. From October 2009 he worked four days a week in this role. His official duties were described as follows:

The Chief Scientific Advisor’s role is to ensure that the Department’s policies and operations, and its contributions to wider Government issues, are underpinned by the best science and engineering advice available.

This new role involved changing his customary garb of shorts and sturdy shoes to a suit and tie as he successfully adapted to the protocols of Whitehall.

He turned out to be extremely effective, using to good effect his ability to understand and engage politicians and civil servants whose backgrounds and skill sets were very different. Those skills had been well honed through years of teaching and public engagement. His advice was highly valued by a succession of ministers, and his impact on the civil servants who worked with him was in many cases transformational. His personal influence on policy was enormous: asking searching questions, questioning assumptions and proposing innovative answers. But he built the capacity of others too, championing and implementing a programme to bring more qualified engineers into the Department and setting a strong lead in working...
closely with the external research community both in academia and in industry. Perhaps his greatest triumph was to convince the government to publish a carbon plan, in December 2011, that drew extensively on his simple but compelling quantitative approach to energy and climate change.

Inspired by the principles embodied in his book on Sustainable energy, he led a team to develop a tool which enabled even the non-expert user to evaluate the consequences of different energy planning strategies. As Mark Lynas has written:

While at DECC, David also launched a 2050 ‘calculator’ that enabled anyone to choose between multiple options for achieving the UK’s mandated 80% carbon emission reductions by 2050. This was then toured around many cities in the UK as a roadshow—I was chair, and David the resident expert. He was witty and confident before audiences—I never saw him stumped by a question, however unlikely. The UK calculator was expanded later into a global calculator, which used the same numbers-based multiple-choice approach to allow anyone to construct their own ‘pathway’ to keeping the planet’s temperature rise below 2C. (Lynas, 2016)

Many other countries took their lead from the 2050 calculator work, which led to several country-specific 2050 calculators. One of the more remarkable outputs of this 2050 global calculator was its indication that, if everyone on Earth became vegetarian, this would mostly solve the global warming problem at a stroke.

Living and working in London and Cambridge did not prevent David from maintaining a phenomenal work-rate, driven by his belief in the importance and urgency of the Department’s task. But David was not a man without hinterland and his life outside work was important too. During his time in DECC he married Ramesh, and their two children, Torrin and Eriska, were born in 2011 and 2014 (see figure 4).
David continued to develop strategies to meet the demands of reducing carbon emissions. In 2012, he pondered how global carbon prices could be successfully negotiated at the COP21 climate conference in Paris 2015. This led to his joint paper proposing a Pledge game, based on game theory, with the rule that all countries that signed up to the game would be legally bound to charge at least the minimum of the carbon prices pledged. The idea was that this would allow countries to express ambitious values with relatively little inhibition and yet to come out with a binding and hopefully non-trivial result.

In 2013, the MacKay–Stone study on potential greenhouse gas emissions associated with shale gas production and use was published by DECC. The overall conclusion was that, subject to suitable safeguards, the effect of shale gas production on UK greenhouse gas emissions would be relatively small. Also in 2013, David published a review article, ‘Could energy-intensive industries be powered by carbon-free electricity?’ Almost his last contribution to the work of DECC was an analysis of the use of biomass as a contribution to meeting the UK’s target for reducing carbon emissions. David’s report on this complex problem, coauthored by Anna Stephenson, was entitled ‘Life cycle impacts of biomass electricity in 2020’, with the subtitle ‘Scenarios for assessing the greenhouse gas impacts and energy input requirements of using North American woody biomass for electricity generation in the UK’. To David’s frustration, this publication was delayed since there were potential conflicts with evolving government policy and it was not necessarily helpful to industry. David felt strongly that the facts should be openly available for discussion.

AFTER DECC

In 2013 David was appointed the first Regius Professor of Engineering in the Department of Engineering at Cambridge, and in the 2016 New Year Honours list he was appointed a Knight Bachelor ‘For Services to Scientific Advice in Government and Science Outreach’.

As his term as Chief Scientific Advisor was drawing to a close, he discussed with Tom Counsell and James Geddes future inference problems which could help to provide better information to the government, in particular to improve understanding of the uncertainties involved in taking different policy decisions. As an example, he discussed the policy on the level of fixed-price payments from small-scale photovoltaics, known as Feed-in-Tariffs. The policy had a complicated set of overlapping objectives, but the existing analysis tools didn’t help understanding of the interaction and trade-offs between these. Even if there had been a single, clear objective for which an optimum policy should be found, the right price to pay for small-scale photovoltaics was uncertain. Such uncertainty was not well treated in the existing predictive packages, and David was disappointed about how hard it was to update the modelling assumptions correctly in the light of new data. His solution was to conceive of a ‘Smart Interpolator to Speed up Complex Models’. As he wrote, this programme would

look over the shoulder of a complex model ... and build an empirical model of the dependence of the outputs of that model on the inputs to the model, which can then serve up rapid answers, predicting the output of the complex model, with error bars indicating the uncertainty of the

---

4 The Sustainable Innovation Forum (SIF15) was the largest business-focused event held during the annual Conference of Parties (COP) on 7–8 December 2015 at COP21 at Stade de France in Paris.
interpolation. The interpolator can also produce derivatives of outputs with respect to inputs, thus enabling rapid optimization of objectives.

Unfortunately, David did not get the time to tackle these challenges before he died.

In May 2015, David began working on information storage using DNA molecules. This followed a 2013 publication in *Nature* by Nick Goldman and his colleagues in which David was acknowledged for his advice on run-length limited (RLL) coding and the information capacity of RLL codes (Goldman et al., 2013). He also advised Goldman and his colleagues in connection with their successful BBSRC grant application on DNA data storage. David and his colleague, Yossy Sayir, worked closely on the kinds of approaches that would be most successful for DNA information storage, based on the likely forms and rates of errors.

David considered this the most ambitious project he had worked on, because there was a strong likelihood that our civilization would proceed to destroy all its knowledge and cultural heritage. Storing it in a safe place encoded in DNA would ensure that future generations, who might one day re-discover how to sequence DNA, could eventually recover the lost information. David identified the main technical hurdle for DNA storage as the disordered nature of the storage medium, where molecules are read in random order. He used methods from information theory to quantify the storage density achievable in DNA, and proposed techniques based on LDPC and fountain coding for practical data encoding in DNA.

### Further Contributions

Alongside the contributions recorded above, David served at various periods on the Royal Society’s International Fellowship Panel, the UKERC Research Committee, the RCUK Energy Programme Scientific Advisory Committee, the World Economic Forum Global Agenda Council on Climate Change, the Advisory Board of Siemens and the Energy Systems Catapult—in the last case, a room has been named after him.

He was politically engaged throughout his adulthood: a member of Cambridge Cycling Campaign, a member of the Campaign for Better Transport (formerly Transport 2000) and Railfuture, a founding member of the Cambridge and St Ives Railway Organization, a supporter of the Campaign against Arms Trade and the Campaign for Nuclear Disarmament, although he was unconvinced by its policy on nuclear power. He was strongly committed to many political issues, which he recorded on his Inference Group website. He campaigned against the Iraq war, for the restoration of Cambridgeshire railways, against miscarriages of justice, the most famous being the Sally Clark case described previously, and, more locally, supporting the staff at Wintercomfort Cambridge who were prosecuted for the drug use of their homeless clients. The last major campaign David was involved in was for safer cycling and walking to and from the North West Cambridge development, despite the gruelling regime of chemotherapy and his illness. He coordinated the campaign with the University, the site team, county council and others, making YouTube videos highlighting the issues, mostly while confined to bed.

In everything he did, David held to and practised high-minded principles. These included making the fruits of his research freely available to all, and living in as eco-friendly a manner as possible. To all his activities, he brought a formidable intellect which gained him worldwide recognition as one of the most inspiring scientists of his generation.
David’s cancer was diagnosed in July 2015 but, undaunted, he was determined to live life to the full, as recorded in his inspiring blog *Everything is connected* (http://itila.blogspot.co.uk/). He underwent chemotherapy, a process he documented on his blog with typical candour, after a careful assessment of whether the likely prolongation of his life would outweigh the fact that more of it would be spent in hospital undergoing treatment. In weighing this judgement, the interests of his very young children, in being able to interact with their father for a few extra months, proved decisive.

David continued to work on DNA storage, but directed much of his energy towards teaching. He continued to teach one graduate course and served as a ‘teaching buddy’ to Jossy Sayir, who had taken over his undergraduate teaching commitments. Despite the heavy burden of cancer treatment, David managed to attend lectures, giving feedback and passing on his approach of carefully choreographed, yet natural, interactive and seemingly improvised lecturing that made him such an exceptional teacher.

In March 2016, just a month before he died, David was present at a wonderful symposium held in his honour, spanning many of his interests—he was very touched by the many tributes paid to his originality and inspiration. David himself gave the final lecture of the symposium, based wholly on mathematical analyses of games that he had played with his children.5

At the funeral of James Clerk Maxwell in 1879, the Reverend Dr Butler, Head Master of Harrow School, paid the following tribute:

> It’s not often, even in this great home of thought and knowledge, that so bright a light is extinguished as that which is now mourned by many illustrious mourners, here chiefly, but also far beyond this place.

The same thoughts were in many minds at David’s well-attended burial, strengthened by the irony that Maxwell had died of the same cause, stomach cancer, and at the same age 48 as David himself. David was buried at the Arbory Trust’s Woodland Burial Site near Barton village outside Cambridge, within view of the radio telescopes of the Radio Astronomy Group with which he had been associated for much of his academic career. The event was, in accordance with his wishes, an occasion for celebration. A procession of family, friends and colleagues was led in bright sunshine to the burial spot by a boisterous yet well-rehearsed Samba band. For many, the abiding memory is that of young children blowing bubbles across and into the grave, in a final act of farewell.

David MacKay is survived by his wife Ramesh Ghiassi MacKay and by their son Torrin and their daughter Eriska.

**Acknowledgements**

We are most grateful to the following for their help in preparing this Memoir: Harry Bhadeshia, Alan Blackwell, Ramesh Ghiassi MacKay, Nick Goldman, John Hopfield, Mark Lynas, Robert MacKay, Radford Neal, Jossy Sayir and Moira Wallace. David provided much invaluable autobiographical material on his website, which can be found at

5 Videos of the two days of lectures can be viewed online at http://divf.eng.cam.ac.uk/djcms2016/.
http://www.inference.phy.cam.ac.uk/mackay/AboutMe.html and which has been extensively used in this Memoir.

HONORS AND AWARDS

2009  Elected Fellow of the Institute of Physics
       Fellow of the Royal Society
2010  Clifford Paterson Lecture, Royal Society
       Elected Fellow of the Institution of Civil Engineers
2012  Honorary Fellowship of the Chartered Institute of Building Services Engineers
2013  Melchett Award, Energy Institute
       Honorary Doctorate, University of Strathclyde
2014  First winner of the Science Communication Medal of the Göttinger Literaturherbstin
2016  Knight Bachelor
       Paradigm Award Winner of the Breakthrough Institute, Oakland, California

REFERENCES TO OTHER AUTHORS


BIBLIOGRAPHY

The following publications are those referred to directly in the text. A full bibliography is available as electronic supplementary material via http://dx.doi.org/10.1098/rsbm.2017.0013 or via https://doi.org/10.6084/m9.figshare.c.3854764.

(2) 1987 A method of increasing the contextual input to adaptive pattern recognition systems, Technical Report RIPR 1000/14/87, Research Initiative in Pattern Recognition. Royal Signals and Radar Establishment, Malvern, Worcester WR14 3PS.


(9) 1994 Bayesian non-linear modelling for the prediction competition. ASHRAE Transactions, 100 Pt 2, 1053–1062. Atlanta, Georgia: ASHRAE.


(15) 1997 (with M. N. Gibbs) Variational Gaussian process classifiers. (https://pdfs.semanticscholar.org/66d4/29b63e0b8e329c565766289b4189c9398174.pdf)


(17) 1998 (With M. C. Davey) Low density parity check codes over GF(q), in Proceedings of the 1998 IEEE Information Theory Workshop, pp. 70–71. IEEE. (doi:0.1109/ITW.1998.706440)


(22) 2002 (With D. J. Ward) Fast hands-free writing by gaze direction. Nature 418, 838. (doi:10.1038/418838a)


(27) 2005 The Trefethen effect. (http://www.inference.phy.cam.ac.uk/mackay/presentations/antiglass/)


(32) 2009  *Sustainable energy: without the hot air*. *CavMag* 1, 7.


