Maximum Entropy in Action

A COLLECTION OF EXPOSITORY ESSAYS

Edited by

BRIAN BUCK
and
VINCENT A. MACAULAY

Maximum Entropy in Action

A COLLECTION OF EXPOSITORY ESSAYS

Edited by

BRIAN BUCK
and
VINCENT A. MACAULAY

Department of Theoretical Physics,
University of Oxford

CLARENDON PRESS · OXFORD
1991

Oxford University Press, Walton Street, Oxford OX2 6DP
Oxford New York Toronto
Delhi Bombay Calcutta Madras Karachi
Petaling Jaya Singapore Hong Kong Tokyo
Nairobi Dar es Salaam Cape Town
Melbourne Auckland
and associated companies in
Berlin Ibadan

Oxford is a trade mark of Oxford University Press

Published in the United States
by Oxford University Press, New York

British Library Cataloguing in Publication Data
A catalogue record for this book
is available from the British Library.

Library of Congress Cataloging in Publication Data
(Data available)
ISBN 0–19–853941–X
ISBN 0–19–853963–0 (pbk)

This book was typeset by the editors in LaTeX on the phototypesetter of the University of London Computer Centre

Printed in Great Britain by Bookcraft (Bath) Ltd
Midsomer Norton, Avon

Preface

This collection of introductory articles on maximum entropy and Bayesian methods grew out of a series of lectures given in the Physics Department of the University of Oxford in the summer of 1989. They were arranged by the editors in association with the Physical Sciences Faculty Board to be interdisciplinary in nature, and drew a large audience from across the spectrum of the sciences.

We were encouraged to try to make them available in printed form and we are grateful to Professors Roger Elliott and Chris Llewellyn Smith for bringing the notion of such a book to the attention of the University Press.

To the editorial team at OUP, who guided us through the new territory of book production with constant patience, many thanks.

The editors 'set the type' themselves from variously word-processed scripts sent by the authors, using Leslie Lamport's user-friendly interface LaTeX to Donald Knuth's TeX program with a 'style' supplied by OUP's house 'TeXnician', for which we thank him. The final high resolution bromides were typeset on the phototypesetter at the University of London Computer Centre by the editors.

The splendid diagrams were mostly prepared in the Drawing Office of the Department of Nuclear and Astrophysics here in Oxford by Irmgard Smith for whose conscientious work on some difficult material we are very grateful. The remainder of the diagrams were supplied by the authors themselves.

Oxford
December 1990

B. B.
V. A. M.

Editors' introduction

Dedicated readers of research journals have probably become aware in recent years of an increasing number of articles mentioning, or even making use of, an intriguingly named new method, and they may have wondered if there was something in it for them. This new technique of maximum entropy is indeed very powerful and has now found application in both practical and theoretical studies ranging from image enhancement to nuclear physics, from statistical mechanics to economics. The reason for this wide application is not that the method is a new physical theory of anything, but that it provides a much needed extension of the established principles of rational inference in the sciences (and possibly elsewhere).

One impediment to the even more general use of such a valuable development is that connected accounts of the method and its implementation are difficult to find in any of the more widely available books and periodicals. It is also regrettably common to overlook mention of the pitfalls. The easily predictable result is that many would-be users have rejected the idea out of hand or have totally misunderstood its purpose. The intention is basically to give a way of extracting the most convincing conclusions implied by given data and any prior knowledge of the circumstances. It is not a magic black box guaranteed to compensate for inadequate data or to rescue badly designed experiments. The method has its roots in probability theory, which has long been recognized as the only consistent way to reason in the face of uncertainty, and it is in fact a modern enrichment of that ancient art of conjecture.

We, the editors of this volume, have been interested for some time in the problem of inversion, of going from incomplete and noisy data to a description of the underlying physical system, and it occurred to us that the principle of maximum entropy might be of help in its solution. In view of the difficulties mentioned above it also seemed to us that it would be useful to arrange a series of personal tutorials on the subject by the active practitioners living in the UK. By disguising these private tutorials as interdisciplinary research seminars, among the first ever to be scheduled at Oxford, we were able to enjoy some excellent instructional sessions. Since the lectures aroused a great deal of interest throughout the University it further occurred to us that published versions of the talks would perhaps stimulate more people to explore the possibilities for themselves.

In this introduction we give a very brief overview of the origins and uses of the idea of maximum entropy in order to set the scene for the detailed expositions to follow. But first we should say that the original choice of the word 'entropy' was probably a mistake, though one that it is now difficult to rectify. The word seems, for no very obvious reason, to inspire deep emotions, especially among those who do not wish the concept of entropy in thermal physics to be identified with the purely mental construct that we are talking about (actually, the two concepts *are* related as we shall see). The word was suggested to Claude Shannon by John von Neumann to denote missing information; but a better choice, and less emotive, would have been the word 'uncertainty'.

Probability theory comes in two parts, one of which is not dealt with at all satisfactorily in the standard texts. The part that everybody agrees with in practice, though not always by using the same justification, has to do with the manipulation and combination of probabilities and the rules are given, essentially completely, by well-known sum and product expressions. The first states that, on any evidence which is not self-contradictory, the probability that some proposition is true, and the probability that the denial of the same assertion is true, add up to unity. The second one tells how to break up the probability of the truth of two propositions, asserted jointly, as the product of the probability that one of them is true, given the other, with the probability that the latter is true. The imposing edifice of the modern theory can be erected on this very slender basis. What is omitted from the rules is how to assign actual numerical values to the probabilities in the first place so that the formalism can lead to useful results. The task of the second part of the theory is to find systematic ways of assigning these numbers, which can include, and go beyond, the usual appeals to statistical information on relative frequencies or to enumeration of possible outcomes which are judged to be equally likely.

At least one component of the missing general method seems to be provided by the *principle of maximum entropy* or, as we would prefer to say, *maximum uncertainty*. The evolution of this idea stems from the work of Claude Shannon in communication technology. He envisaged that it should be possible to attach to a probability distribution a single number measuring the total amount of uncertainty represented by that distribution. In short, we would feel less uncertain about the real state of affairs, if only a few possibilities out of many were at all likely, than if we had to take seriously a lot more of the possibilities, each with some appreciable probability of being true. By using an argument of consistency (two routes to a measure should give the same answer), a requirement of continuity and a plausible rule for the combination of uncertainties, Shannon arrived at an effectively unique expression for the desired measure, at least for the commonly considered case of propositions which are exhaustive and mutually exclusive in a given situation. These last conditions entail that only one of the proposed assertions is true, though of course we are uncertain which one. This is why we need to use probability theory. The result can be generalized in various directions, for example, in order to apply it to exhaustive but non-exclusive propositions, but we shall consider here a different line of development.

In the simplest case, we have N propositions labelled by $n = 1, 2, \ldots, N$, exactly one of which is true, and by some means we have assigned to them a set of probabilities p_n which encode all our knowledge relevant to choosing one rather than another. The amount of uncertainty in this situation is then, according to Shannon, represented by the formula

$$S = -k \sum_{n=1}^{N} p_n \ln p_n,$$

in which k is an arbitrary positive constant. With suitable choice of k, the result can be interpreted as an estimate of the number of questions (having yes/no answers) which would be needed to isolate the true proposition. In particular, most questions would be needed when all $p_n = 1/N$, i.e., for the uniform distribution of probabilities. Hence the result does seem a reasonable embodiment of the qualitative idea of uncertainty. This expression for the 'entropy' of a probability distribution was used by Shannon in the proof of his fundamental theorem on the efficient coding of messages for transmission over noisy communication channels. The surprising conclusion is that transmission can be made essentially error-free.

The above formula had also appeared long before in physics. It is identical in form to Planck's expression for thermodynamic entropy, in which the probabilities refer to the possible occurrence of various microstates of a macroscopic body and, of course, the quantity S is taken to be a down-to-earth physical property of a body, not a mere summary of our knowledge or uncertainty about the state of the system. Still, it was striking that exactly the same mathematical expression had appeared in a completely different context and it should perhaps have signalled immediately that there was some concept in common. The great contribution of E. T. Jaynes, some years later, was to point out the underlying connection. His idea was that in both examples the formula did indeed reflect in a quantitative way the total amount of uncertainty remaining after all relevant information about some situation had been taken into account. Even more importantly, he proposed that it could be made the basis for a new method of assigning probabilities.

Jaynes' principle, as it is now called, uses *testable* information, that is, propositions which are relevant to a probability distribution and whose truth can be checked when that distribution has been assigned. An example is data on the value of a measured quantity which could be taken as an average over a probability distribution of various outcomes. The principle then states that the best choice of probabilities is that for which the uncertainty number is maximized, subject to the constraints implied by the testable information. Maximizing uncertainty, while ensuring that known data are reproduced, definitely corresponds well with intuitive ideas of honest probability assignment. This variational principle has several virtues, not the least of which is that for common types of testable information it can be proved that the maximizing distribution is unique. No possibility is assigned zero probability unless the data explicitly require it and in fact the resulting distribution is as spread out among the possibilities as is compatible with the known constraints. If, indeed, there is no information other than that the propositions are exhaustive and mutually exclusive (so that the probabilities sum to unity), then the principle of maximum uncertainty yields the uniform distribution, which is in pleasing agreement with common sense.

The expression for uncertainty is easily extended to cover a countable infinity of possibilities and with this form Jaynes showed that his principle gives a convincing basis for the canonical and grand canonical distributions of statistical mechanics. These refer to the probabilities of occurrence of energy eigenstates of a physical system at equilibrium and their assignment from information theory assuming fixed mean energy and particle number shows very clearly that their form does not depend on physics, but rather that they represent our knowledge of the system. Even more striking is the result that if k is chosen as the Boltzmann constant then the thermodynamic entropy and the maximized uncertainty function are numerically equal. Many features of thermodynamics are then seen in an entirely new light and the way is clear to lay solid foundations for a theory of non-equilibrium processes. These results are far-reaching and illuminating, but the real point of the work of Jaynes is the new method of inference, with its numerous potential applications in other fields.

For many such applications the uncertainty formula needs to be modified so that assignment of continuous probability density distributions can be handled. It has now been shown quite convincingly by several authors that the correct form for the uncertainty of a continuous distribution $p(x)$ is

$$S = -\int p(x) \ln \left(\frac{p(x)}{m(x)} \right) \mathrm{d}x,$$

where $m(x)$ is a function determined by the exact nature of the problem. Maximization of this uncertainty shows, for example, that when it is sensible to choose $m(x)$ a constant, and the mean value and variance of a quantity are specified, then the maximally non-committal density is the normal or Gaussian distribution. A further natural generalization is to probabilities in function spaces. Such an extension will always be needed if we are to have a proper rationale for tackling inverse problems. For it is required to construct, from noisy and discrete data and background information on the signal, the probabilities of different functional forms for that signal. The finding of effective ways of applying the principle of maximum uncertainty in function spaces is a topic of current research.

Pending a natural solution of inverse problems along the lines sketched out above, there has evolved an alternative scheme for the reconstruction of signal sources which take the form of positive density distributions. Examples are light intensities of images, particle number densities and spectra of many kinds, all viewed through some recording instrument which may change the original signal to another physical form, discretize it and simultaneously introduce error. This alternative algorithm was pioneered by Drs Gull and Daniell and is also known as the *method of maximum entropy*, but it is not an obvious continuation of the ideas outlined so far. It is still a variational principle and it is almost certainly related to the earlier method, though the connections are not yet completely elucidated.

The main strategy is to imagine the building up of the investigated positive density by placing numerous small quanta into a finite number of cells, so that the density distribution is reasonably well approximated by specifying the numbers of quanta in the various boxes. It is then assumed that the best distribution is that which can be made by the above process in the greatest number of ways while still agreeing with the known data according to some criterion. Thus the method should perhaps be called the *principle of greatest multiplicity* rather than of *maximum entropy*. The final form of the procedure involves the maximization, under data constraints, of the so-called *configurational entropy* of the density distribution. This latter quantity has the same form as the information theory entropy, but expressed in terms of the proportions of the quanta in the cells rather than involving probabilities. Hence the actual procedure followed looks very much like the generation of a probability distribution by means of Jaynes' principle, though the object produced is a physical density function. The method shares some of the intuitively desirable properties of the idea of maximizing uncertainty. In particular, the deduced density is as spread out or uniform as is possible while remaining compatible with the data and it does not contain any feature for which there is no evidence in the data. Furthermore, Dr Skilling has shown, using reasonable axioms, that the positive density of greatest multiplicity is also the most probable one among all those which agree with observation.

The various ideas discussed here are clearly interrelated and we expect that future researches will converge on some generally acceptable philosophy for attacking the difficult and inescapable problems of probabilistic inference. This book gives examples of the present state of the art and is organized as follows. The first two chapters describe the rationale of the maximum entropy method, while Chapters 3–5 explore some applications. The next two chapters then make plain the interpretation of thermodynamic entropy in terms of maximized uncertainty, and the relation of both to the Bayesian probability theory. Finally, an alternative view of these things is given in the context of crystallography.

In more detail, *Of maps and monkeys* gives an overview of the different kinds of tasks that a unified approach to data handling is required to deal with. The only consistent calculus for this process of *inference* is that of probability theory as championed by Laplace, Jeffreys, Cox and Jaynes. Dr Daniell demonstrates that in many problems the amount of data available is extremely small in comparison with the 'size' of the image we are trying to reconstruct from it. As a result, it is not satisfactory to ignore the variation over different images of the *prior probability distribution*—an encoding of what is known about the real image before we consider the current data—and to rely only on the information carried by the data (via the *likelihood*). The intuitively appealing *monkey argument* is used to generate the entropic prior on positive, additive images, a prior which favours reconstructions as uniform as possible. Mathematical complexity is avoided and the informal style of the original talk has been deliberately retained.

In *Fundamentals of MaxEnt in data analysis*, a more sophisticated argument is supplied to justify the status of the entropic prior. This is important: for although the results provided by using MaxEnt speak for themselves, as will be seen later, it is central to the Bayesian outlook that methods follow from plausible axioms in a logically correct fashion; we want to eliminate *adhoc*kery from our procedures. Central to this new derivation is a requirement that our inferences should depend only on objects which Dr Skilling has called *observables*: integrals linear in the image. This is not an easy chapter, but out of the mathematics comes an extension of MaxEnt: the ability to introduce a *preblur*. The traditional entropy prior assumes no correlations between the pixels of an image. However this has never really been satisfactory since images are in fact almost always correlated. The new formalism allows correlations to develop in the image if the data contains evidence for them. This is an exciting new area and we anticipate that future developments of these methods will involve the incorporation of more specific accumulated experience about the nature of particular sorts of images.

The effort in the applications chapters has been to give enough details of the particular field so that the problems to which MaxEnt has been applied are clear in principle. *Maximum entropy and nuclear magnetic resonance* gives a brief account of modern time-domain NMR before considering in detail the pros and cons of using MaxEnt to recover spectra. The case for the use of MaxEnt in this sort of spectroscopy is still somewhat controversial. In general, the spectrum is not a positive, additive distribution: indeed it is not even real. So the last word has not yet been said on Bayesian techniques for the analysis of such signals. Dr Hore presents an alternative formulation, due to Dr Daniell and himself, more in the spirit of the statistical mechanics which was the origin of the use of entropy in inference. Their entropy is defined on a quantum-mechanical density operator and is a genuine encoding of uncertainty.

In Chapter 4, we again look at spectroscopy, this time Raman as well as NMR. Some impressive reconstructions are displayed, which should convince people that there is something important going on here. In a final section, the disaster which results from using *ad hoc* forms of 'entropy' is displayed.

In *Maximum entropy and plasma physics*, we see MaxEnt being used routinely and skilfully as part of the toolkit of a physicist confronted with a variety of the diagnostic problems which arise in the study of confined plasmas. Here the problems of noisy and sparse data can be quite acute. A useful technique for tuning the values of imperfectly known parameters is described.

At this point we make a couple of asides, which the neophyte might wish to skip until he is more familiar with the material. Firstly, we remark on a comment which is common to several of the chapters saying something to the effect that certain physical distributions that are positive (be they of charge, mass, spectral intensity or whatever) 'can be regarded as probability distributions'. As we have said earlier in this introduction, it appears to us that this remark is rather confusing. It seems to be made only to motivate the use of the Gibbs/Shannon/Jaynes entropy as a measure of the uniformity (or information content) of any physical distribution, which it might seem appropriate to maximize (or minimize). However, the real justification comes from arguments such as those of Chapter 2 where certain assumptions on the structure of images lead naturally to the *configurational* entropy reflecting the prior probability of an image. The message is that the entropies of positive images and probability distributions are rather different beasts.

Secondly, the majority of applications of MaxEnt in this book use the constraint $\chi^2 = N$ to incorporate the data. This is a somewhat *ad hoc* technique, justified by the fact that, on average, every data point is one standard deviation away from its true value. Thus it is a 'long run' or *frequentist* rule. In practice what this constraint does is to set the relative weight given to the entropy and to the data (via χ^2) in determining the resultant image: it balances uniformity against (possibly spurious) structure. An alternative and Bayesian way to proceed is to enlarge the *hypothesis space* to include the relative weighting of these terms among the parameters to be estimated. This is the role played by α in Chapter 2. It will in general lead to a χ^2 not equal to N.

Macroirreversibility and microreversibility reconciled describes how the conceptual difficulties of the second law of thermodynamics disappear when the methods of statistical physics are recognized as instances of reasoning from incomplete information. The second law describes the loss of information which results when a system is perturbed and note is taken only of its final equilibrium state, not of its intermediate dynamical evolution. The thermodynamic entropy should be identified with the maximized Gibbs/Shannon/Jaynes entropy. The ideas are then generalized to set up a formalism for non-equilibrium processes. In fact this chapter is a concentrated crash-course in Bayesian techniques and its arguments are quite general. In three appendices, further side-issues are addressed. Again a certain mathematical sophistication is assumed of the reader.

In *Some misconceptions about entropy*, Dr Garrett's arguments of the previous chapter are amplified and illustrated. The Boltzmann H function, an early candidate for the statistical representation of thermodynamic entropy, is shown to be wanting, failing to describe other than non-interacting systems. Dr Gull demonstrates in a new way how Brownian motion, an irreversible process with which orthodox statistical physics has had some difficulty, can be accounted for. This can be done by the introduction of a maximum entropy probability distribution to describe the space–time trajectories of particles, with constraints coming from the dynamics of the process. This chapter closely follows the original talk and some of the derivations are rather compressed.

The final chapter, *The X-ray crystallographic phase problem*, presents a brief review of current ideas on how to tackle the extremely demanding inverse problem that arises in the analysis of X-ray diffraction patterns. There emerges an alternative derivation of the entropic prior which is couched to some extent in the language of orthodox statistics. It is useful to have different derivations of central results and Dr Bricogne's arguments, using as they do the saddlepoint approximation, recall the famous 'justification' of the maximum entropy distribution which goes under the name of the Darwin–Fowler method.

About the authors

Babul Baruya

BP Research Centre
Chertsey Road
Sunbury-on-Thames
Middlesex, TW16 7LN

A. Baruya is a senior statistician at BP Sunbury. He specializes in the development and application of advanced statistical and data processing techniques.

Gérard Bricogne

MRC Laboratory of Molecular Biology
Hills Road
Cambridge, CB2 2QH
and
LURE
Bâtiment 209d
91405 Orsay
France

Dr G. Bricogne was born and educated in France, where he graduated in mathematics and chemistry. He holds a Ph.D. from the University of Cambridge, for research done at the MRC Laboratory of Molecular Biology on phase determination methods based on the exploitation of non-crystallographic symmetry. The computational aspects of this work led to the first determinations of virus structures to atomic resolution. His subsequent research has been devoted to extending 'direct' methods of X-ray crystal structure analysis, well known for small molecules, to macromolecules. His interests lie in the methods of molecular structure determination and representation, and especially in the underlying mathematics.

Dr Bricogne has held positions as a research fellow of Trinity College, Cambridge (1975–81), as an assistant professor of biochemistry at the College of Physicians and Surgeons of Columbia University, New York (1981–83), and as a director of research at the French Synchrotron Radiation Facility (LURE) in Orsay (1983–present). Since 1988 he has been working as a visiting scientist at the MRC Laboratory of Molecular Biology and at Trinity College, Cambridge.

Geoff Cottrell

JET Joint Undertaking
Abingdon
Oxfordshire, OX14 3EA

G. A. Cottrell is principal scientific officer in the physics group (radio frequency heating) at the Joint European Torus (JET) experiment based at the UKAEA's Culham Laboratory in Oxfordshire, UK. After reading physics at Sussex University, he received his Ph.D. in radio astronomy at the University of Cambridge in 1977. He spent two years in postdoctoral research in low temperature solid state physics at UMIST, before transferring to Culham Laboratory in 1979 to work on controlled thermonuclear fusion. He joined the JET project full-time in 1985. Current areas of study include the interpretation of plasma physics experiments on energy confinement and heating as well as the development of a physical model capable of both explaining present experimental results and predicting future tokamak performance. He is currently a research fellow at Wolfson College, Oxford.

Geoff Daniell

Department of Physics
The University
Southampton, SO9 5NH

G. J. Daniell read natural sciences at Downing College, Cambridge and went on to gain a Ph.D. at the Cavendish Laboratory, on the theory of a radio aerial immersed in the ionosphere. He is currently a senior lecturer in the Department of Physics at the University of Southampton. His research interests include data and signal processing techniques, maximum entropy and the foundations of probability theory and computational techniques in physics with especial emphasis on problems in quantum field theory and statistical mechanics.

Simon Davies

BP Research Centre
Chertsey Road
Sunbury-on-Thames
Middlesex, TW16 7LN

S. Davies joined BP Research in 1988 after completing his doctorate in nuclear magnetic resonance (NMR) spectroscopy at Oxford University. His interests include frequency-selective excitation, the use of relaxation data to probe the internal structure of porous media, NMR imaging and inverse problems.

Anton Garrett

Department of Physics and Astronomy
University of Glasgow
Glasgow, G12 8QQ

A. J. M. Garrett took his first degree in physics at the University of Cambridge. He remained there and gained his Ph.D. in kinetic theory in 1984. It was while he was a research fellow at Magdalene College that his conversion to the Bayesian outlook occurred, through his contact with Steve Gull. From 1985–88, he held a research fellowship at the University of Sydney and currently he is a Royal Society of Edinburgh research fellow at the University of Glasgow.

His interests are wide-ranging with most of his recent work centred on quantum philosophy and the foundations and applications of probability theory. On the latter, he is at present writing a book, to be called *'Inference and Inferential Physics'*.

He is a member of UK Skeptics and is a keen debunker of pseudoscience.

Andrew Grant

BP Research Centre
Chertsey Road
Sunbury-on-Thames
Middlesex, TW16 7LN

A. Grant joined BP Research in 1984 after completing his doctorate in time-resolved, laser flash photolysis electron spin resonance at Oxford

University. His interests include the application of appropriate laser and optical spectroscopies to industrial problems, and the recovery of enhanced information from spectroscopic data, recorded under industrially relevant conditions.

Steve Gull

Cavendish Laboratory
Madingley Road
Cambridge, CB3 0HE

S. F. Gull took his first degree in theoretical physics at the University of Cambridge. His Ph.D. was earned in the radio astronomy group at the Cavendish Laboratory there. He has remained a member of St John's College and is now a university lecturer in physics.

Dr Gull has done work in the dynamics of radio galaxies and supernovae, the astronomy of the cosmic background and numerical hydrodynamics and plasma physics. He was one of the leading figures in the establishment of the maximum entropy method applied to image enhancement and now works on general maximum entropy data processing, the foundations of probability theory and inverse problems of all types. Currently he is working extensively on the physical applications of Clifford algebras.

With John Skilling, he co-founded Maximum Entropy Data Consultants Ltd.

Peter Hore

Physical Chemistry Laboratory
South Parks Road
Oxford, OX1 3QZ

P. J. Hore is a university lecturer in Physical Chemistry and fellow of Corpus Christi College at the University of Oxford. Having got both his degrees from Oxford, he spent two years as a Royal Society European programme research fellow with Professor R. Kaptein at the University of Groningen. He returned from the Netherlands to a junior research fellowship at St John's College, Oxford to work with Dr (now Professor) R. Freeman, and took up his present position in 1983. His research interests include energy conversion in photosynthetic reaction centres, spin effects in chemical reactions, protein structure and folding, and the development of new methods in nuclear magnetic resonance spectroscopy.

Ken Packer

BP Research Centre
Chertsey Road
Sunbury-on-Thames
Middlesex, TW16 7LN

Professor K. Packer is a chief research associate with BP Research at their Sunbury-on-Thames (UK) research centre. He manages the spectroscopy, microscopy and structural crystallography activities within the Analytical Support and Research Division. Prior to joining BP Research in 1984, he held a personal chair in the School of Chemical Sciences, University of East Anglia, Norwich. His research interests centre on nuclear magnetic resonance and its application to a wide range of physico-chemical areas.

John Skilling

Department of Applied Mathematics and Theoretical Physics
Silver Street
Cambridge, CB3 9EW

J. Skilling was an undergraduate at Cambridge, obtaining his first degree in physics in 1965, following which he joined the radio astronomy group in the Cavendish Laboratory for his doctoral work on plasma instabilities in astrophysics. He spent two years abroad at Princeton University (1969–71) working at the Plasma Physics Laboratory, and then returned to Cambridge to join the teaching staff of the Department of Applied Mathematics and Theoretical Physics, with cosmic ray physics as his major research interest.

In 1977, his research career changed abruptly through the seminal work on maximum entropy image reconstruction by Gull and Daniell. Since then, he has concentrated on this field, contributing to both theory and practice. In 1980, he founded Maximum Entropy Data Consultants Ltd with Dr Gull, to assist in the development of a wide variety of practical industrial and academic applications of maximum entropy techniques.