Preface

This book is the outcome of a workshop entitled "*Computational information geometry for image and signal processing*" (http://www.icms.org.uk/workshops/infogeom) that was hosted by the International Centre for Mathematical Sciences (ICMS, Edinburgh, UK), with funding from the London Mathematical Society and the British Engineering and Physical Sciences Research Council. The workshop took place during September 21–25, 2015. Participants at this workshop (see group photo of Fig. 1) were kindly asked to submit for peer-review a chapter summarizing some of their recent research achievements in this area.

First, let us give some background on the workshop. The workshop began with lectures providing overviews of information geometry and its applications, computational aspects, and in particular applications in machine learning, cognition, medical imaging of the brain and radar processing. The common ground is a scientific context in which the experimenter has a model with, possibly intricately connected, statistical properties. Then, the desired identification and extraction of features of interest depends on an optimisation process that exploits the existence of a metric structure on smooth families of probability density functions, or approximations thereto from observed frequency data for example, typically using relative information entropy. This was followed by a session on cases when the statistical aspects could be represented *a priori* by mixtures of given probability density functions notably multivariate Gaussians where, typically, the analytic information metric is not known and approximations need to be made. One of the challenges in computational anatomy is the identification of anomalous shape features. In this context deformed exponential families of probability density functions, originally introduced in the study of statistical physics of strongly correlated systems, can be used to select a model for statistical shape via a segmentation of the target organ region and suitable training data. A number of presentations were concerned with various aspects of nonlinear filters where a signal process is progressively estimated from the history of a related process of observations. Information theoretic properties of filters have relevance in non-equilibrium statistical mechanics. The nonlinear filtering situation is typically infinite dimensional and the role of finite-dimensional approximations is an important research area. For example, the

Fig. 1 Group picture (photo courtesy of ICMS, UK)

manifold projection method for stochastic nonlinear filtering gives the projection filter. Using a Hilbert space structure on the space of probability densities, the infinite-dimensional stochastic partial differential equation for the optimal filter can be projected onto a finite dimensional exponential or mixture family, respectively, with two different metrics: the Hellinger distance and the L_2 direct metric. Such projection filters can match the performance of numerical methods based on much higher numbers of parameters, thereby providing both estimation and computational efficiency. Hamilton Monte Carlo methods have an important role in generic Bayesian inference problems and presentations emphasised its geometric foundation, and how they can be used to understand properties like ergodicity and guide the choice of tuning parameters in applications. The formal foundations of the Hamilton Monte Carlo algorithm were examined through the construction of measures on smooth manifolds to demonstrate an efficient and robust implementation.

The balance of presentations and a novel departure including substantial discussion sessions with lively exchanges seems to have proved correct to judge from the very positive feedback. Evidently this balance achieved some of the aims of the workshop since the feedback reported the transfer of new skills and new ideas generating new approaches and applications. Moreover, new and renewed contacts were reported and expected to lead to prospective collaborations.

The workshop gave us an opportunity to discuss further research that we overview concisely below. The final session drew together themes from the lengthy, thought-provoking and, ultimately, very fruitful discussion sessions that were held throughout the workshop. A number of particularly promising areas for future work were highlighted:

- Dimension reduction: generically important, a strategic emphasis for future work is the infinite to finite case; this builds upon and develops recent successes in nonlinear filtering.

- Simplex approach: a variety of contributors, from a variety of perspectives, pointed to the great challenges and, correspondingly, large potential benefits associated with a move from traditional manifold-based methodologies to their simplicial counterparts; these being suitable closures, there are strong connections to ongoing parallel work on the closure of exponential families.
- Geometry of model—dually, data—space: this was a recurring theme throughout the Workshop and, especially, its Discussion sessions; among other novel possibilities, the potential benefit of a twisted product of Wasserstein and Fisher–Rao geometries was noted.
- Computational Information Geometry: its continuing evolution in three streams, and their potential fruitful interactions, were evident throughout the Workshop; in brief:

  – Statistical perspective:

    Offers an operational universal model space modulo, where appropriate, (coarse) binning;
    delivers novel, unique contributions to ubiquitous, challenging inference problems including: handling model uncertainty, estimating mixture models and big discrete data;
    overall motivation here is the long-term aim of providing experimenters with new operational tools for handling their data.

  – Engineering perspective: The workshop provided a unique venue for discussing fruitfully more precisely what is meant by "Computational Information Geometry", especially with respect to the established field of "Information Geometry". Several strong points have been acknowledged and promising research avenues sketched. For example:

    Since we deal with large but finite data sets, instead of considering parametric probability families (IG), consider dealing with high-dimensional data simplices with potentially many empty bins (therefore technically different from the regular probability simplex of IG). This raises the theoretical issue of discrepancy between dealing with continuous versus discrete models, etc.
    Some problems change fundamentally from the computational point of view if we consider the standard reduction by sufficient statistics (IG): See for example, Montanari, A. Computational implications of reducing data to sufficient statistics. Technical Report 2014-12, Stanford University, 2014.
    The number of local extrema in deep learning networks and how to analyse computationally the set of singularities. The workshop had about one third of papers dealing with implementations for applications ranging from signal processing, to medical imaging, to clustering, and to statistical mixture estimation, among others.

  – Artificial Intelligence and Machine Learning: Papers included several aspects of machine learning and optimisation thereof. General parametrised geometric objects allow design of efficient learning systems by imposing natural geometric constraints. We learnt of promising possibilities for spontaneous data learning introducing a novel explanatory paradigm beyond the discussion for misspecification of a parametric model. A recent breakthrough in the computation of optimal transport barycenters has the potential to impact deeply the machine learning and imaging communities.

- Singularities and unboundedness: once again, this new theme emerged during the Workshop, differing emphases being to the fore in different presentations; unifying these, and of central importance, is the fundamental move from open to closed structures, and consequences thereof.
- Divergences: the fundamental links between these naturally asymmetric objects and information geometry and, especially, their strategic actual or potential benefits in application were underscored by vital new contributions made during the meeting.
- Optimisation: a variety of opportunities for new or improved optimisation methodologies were opened up by advances reported during the Workshop notably, those based on natural gradient methods and still others arising in the engineering strand of computational information geometry.
- Markov Chain Monte Carlo: Hamiltonian geometry: distinctive results and methodologies, summarised above, reported during the Workshop open up a wide range of new potential applications, especially in Bayesian statistics; intriguingly, it will be of great interest to see how far links can be established between typical sets, support sets and structural zeroes.

We thank all the invited participants at this workshop:

Amari, Shun-ichi	RIKEN Brain Science Institute
Anaya-Izquierdo, Karim	University of Bath
Armstrong, John	King's College London
Ay, Nihat	Max Planck Institute for Mathematics in the Sciences
Barbaresco, Fréderic	Thales Land and Air Systems
Belavkin, Roman	Middlesex University
Betancourt, Michael	University of Warwick
Brigo, Damiano	Imperial College London
Byrne, Simon	University College London
Critchley, Frank	the Open University
Dodson, Kit	University of Manchester
Eguchi, Shinto	Institute of Statistical Mathematics
Galanis, George	Hellenic Naval Academy
Goh, Alvina	National University of Singapore
Jupp, Peter Edmund	University of St Andrews
Komori, Osamu	the Institute of Statistical Mathematics
Marti, Gautier	École Polytechnique & Hellebore Capital Management
Matsuzoe, Hiroshi	Nagoya Institute of Technology
Matúš, František	Institute of Information Theory and Automation
Newton, Nigel	University of Essex

Nielsen, Frank	École Polytechnique & Sony Computer Science Laboratories
Nock, Richard	NICTA & the Australian National University
Ohara, Atsumi	University of Fukui
Perrone, Paolo	Max Planck Institute for Mathematics in the Sciences
Peter, Adrian M.	Florida Institute of Technology
Peyré, Gabriel	Université Paris-Dauphine
Pistone, Giovanni	Collegio Carlo Alberto
Sabolova, Radka	the Open University
Saint-Jean, Christophe	Université de La Rochelle—UFR Sciences
Sampson, W.	University of Manchester
Schwander, Olivier	LIP6, UPMC
Szkola, Arleta	Max Planck Institute for Mathematics in the Sciences
Takatsu, Asuka	Tokyo Metropolitan University
Van Bever	Germain, the Open University
Vos, Paul	East Carolina University
Zhang, Jun	the University of Michigan at Ann Arbor

This book is a collection of eleven chapters that span both the theoretical side and practical applications of computational information geometry for image and signal processing:

1. Information Geometry and Its Applications: *An* Overview by Frank Critchley and Paul Marriott
2. The Geometry of Model Sensitivity: An Illustration by Karim Anaya-Izquierdo, Frank Critchley, Paul Marriott, and Paul Vos
3. On the Geometric Interplay Between Goodness-of-Fit and Estimation: Illustrative Examples by Karim Anaya-Izquierdo, Frank Critchley, Paul Marriott and Paul Vos
4. Spontaneous Learning for Data Distributions via Minimum Divergence by Shinto Eguchi, Akifumi Notsu, and Osamu Komori
5. Extrinsic Projection of Itô SDEs on Submanifolds with Applications to Non-linear Filtering by John Armstrong and Damiano Brigo
6. Fast $(1 + \epsilon)$ -Approximation of the Löwner Extremal Matrices of High-Dimensional Symmetric Matrices by Frank Nielsen and Richard Nock
7. Dimensionality Reduction for Information Geometric Characterization of Surface Topographies by C.T.J. Dodson, M. Mettaneny and W.W. Sampson
8. On Clustering Financial Time Series: A Need for Distances Between Dependent Random Variables by Gautier Marti, Frank Nielsen, Philippe Donnat, and Sébastien Andler
9. The Geometry of Orthogonal-Series, Square-Root Density Estimators: Applications in Computer Vision and Model Selection by Adrian M. Peter, Anand Rangarajan and Mark Moyou
10. Dimensionality Reduction for Measure Valued Evolution Equations in Statistical Manifolds by Damiano Brigo and Giovanni Pistone

11. Batch and Online Mixture Learning: A Review with Extensions by Christophe Saint-Jean and Frank Nielsen

We express our gratitude to the peer reviewers for their careful feedback, which led to the polished final form of these chapters.

Throughout the planning period, and during the running of the Workshop, the ICMS staff provided excellent support to the organisers. The facilities and catering were excellent. We mention in particular Moira Spencer for her tireless attention to detail and friendly unflappable handling of inevitable complications arising from a large number of international participants.

There is an exciting time ahead for computational information geometry, studying further the fundamental concepts and relationships of information, geometry and computation, and we envision many more applications in signal and image processing.

Palaiseau, France	Frank Nielsen
Milton Keynes, UK	Frank Critchley
Manchester, UK	Christopher T.J. Dodson
June 2016	

http://www.springer.com/978-3-319-47056-6

Computational Information Geometry
For Image and Signal Processing
Nielsen, F.; Critchley, F.; Dodson, C.T.J. (Eds.)
2017, XIV, 299 p. 79 illus., Hardcover
ISBN: 978-3-319-47056-6