Statistical Challenges in 21st Century Cosmology
Proceedings IAU Symposium No. 306, 2014
A. F. Heavens, J.-L. Starck & A. Krone-Martins, eds.

doi:10.1017/S1743921314011120

Bayesian large-scale structure inference: initial conditions and the cosmic web

Florent Leclercq[1,2,3] **and Benjamin Wandelt**[1,2,4]

[1]Institut d'Astrophysique de Paris (IAP), UMR 7095, CNRS - UPMC Université Paris 6, 98bis boulevard Arago, F-75014 Paris, France

[2]Institut Lagrange de Paris (ILP), Sorbonne Universités, 98bis boulevard Arago, F-75014 Paris, France

[3]École polytechnique ParisTech, Route de Saclay, F-91128 Palaiseau, France

[4]Departments of Physics and Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801

emails: `florent.leclercq@polytechnique.org`, `wandelt@iap.fr`

Abstract. We describe an innovative statistical approach for the *ab initio* simultaneous analysis of the formation history and morphology of the large-scale structure of the inhomogeneous Universe. Our algorithm explores the joint posterior distribution of the many millions of parameters involved via efficient Hamiltonian Markov Chain Monte Carlo sampling. We describe its application to the Sloan Digital Sky Survey data release 7 and an additional non-linear filtering step. We illustrate the use of our findings for cosmic web analysis: identification of structures via tidal shear analysis and inference of dark matter voids.

Keywords. large-scale structure of universe, methods: statistical

1. Introduction

How did the Universe begin? How do we understand the shape of the present-day cosmic web? Within standard cosmology, we have an observationally well-supported model for the initial conditions (ICs) – a Gaussian random field – and the evolution and growth of cosmic structures is well-understood in principle. It is therefore natural to analyze large-scale structure (LSS) surveys in terms of the simultaneous constraints they place on the statistical properties of the initial conditions of the Universe and on the shape of the cosmic web. Due to the computational challenge and to the lack of detailed physical understanding of the non-Gaussian and non-linear processes that link galaxy formation to the large-scale dark matter distribution, the current state of the art of statistical analyses of LSS surveys is far from this ideal and these problems are addressed in isolation. Here, we describe progress towards the full reconstruction of four-dimensional states and illustrate the use of these results for cosmic web classification.

2. Statistical approach

2.1. *Why Bayesian inference?*

Cosmological observations are subject to a variety of intrinsic and experimental uncertainties (incomplete observations – survey geometry and selection effects –, cosmic variance, noise, biases, systematic effects), which make the inference of signals a fundamentally ill-posed problem. For this reason, no unique recovery of the initial conditions and of the shape of the present-day cosmic web is possible; it is more relevant to quantify a probability distribution for such signals, given the observations. Adopting this point of view for large-scale structure surveys, Bayesian forward modeling (gravitational structure formation is the generative model for the complex final state, starting from a simple initial state – Gaussian or nearly-Gaussian ICs) offers a conceptual basis for dealing with the problem of inference in presence of uncertainty (e.g. Jasche & Wandelt 2013a; Kitaura 2013; Wang *et al.* 2013).

2.2. *High-dimensionality*

Statistical analysis of LSS surveys requires to go from the few parameters describing the homogeneous Universe to a point-by-point characterization of the inhomogeneous Universe. The latter description typically involves tens of millions of parameters: the density in each voxel of the survey volume. No obvious reduction of the problem size exists. "Curse of dimensionality" phenomena (Bellman 1961) are therefore the significant obstacle in this high-dimensional data analysis problem. They refer to the problems caused by the exponential increase in volume associated with adding extra dimensions to a mathematical space, and therefore in sparsity given a fixed amount of sampling points. Numerical representations of high-dimensional probability distribution functions (pdfs) will tend to have very peaked features and narrow support, which means that traditional sampling methods will fail. However, gradients of these functions carry capital information, as they indicate the direction to high-density regions, permitting fast travel through a very large volume in parameter space.

2.3. *Hamiltonian Monte Carlo*

The Hamiltonian Monte Carlo algorithm (Duane *et al.* 1987) is an algorithm for exploring parameter spaces with particles (samples). The general idea is to use classical mechanics to solve statistical problems. The algorithm interprets the negative logarithm of the pdf to sample from, $\mathcal{P}(\mathbf{x})$, as a potential, $\psi \equiv -\ln(\mathcal{P}(\mathbf{x}))$, and integrates Hamilton's equation in parameter space. Due to the conservation of energy in classical mechanics, the theoretical acceptance rate is always unity. Therefore, HMC beats the "curse of dimensionality" by exploiting gradients ($\partial\psi(\mathbf{x})/\partial\mathbf{x}$ in Hamilton's equations) and using conserved quantities.

3. Physical reconstructions

3.1. *Bayesian large-scale structure inference in the SDSS DR7*

The full-scale Bayesian inference code BORG (Bayesian Origin Reconstruction from Galaxies, Jasche & Wandelt 2013a) uses HMC for four-dimensional inference of density fields in the linear and mildly non-linear regime. The (approximate) physical model for gravitational dynamics included in the likelihood is second-order Lagrangian perturbation theory (2LPT), linking initial density fields (at a scale factor $a = 10^{-3}$) to the presently observed large-scale structure (at $a = 1$). The galaxy distribution is modeled as a Poisson sample from these evolved density fields. The algorithm also accounts for luminosity dependent galaxy biases (Jasche & Wandelt 2013b). In Jasche *et al.* (2015), we apply the BORG code to 463,230 galaxies from the `Sample dr72` of the New York University Value Added Catalogue (NYU-VAGC, Blanton *et al.* 2005), based ot the final data release (DR7) of the Sloan Digital Sky Survey (SDSS, Adelman-McCarthy *et al.* 2008; Padmanabhan *et al.* 2008).

Each inferred sample (Fig. 1, left) is a "possible version of the truth" in the form of a full physical realization of dark matter particles. The variation between samples (Fig. 1, right) quantifies joint and correlated uncertainties inherent to any cosmological observation and accounts for all non-linearities and non-Gaussianities involved.

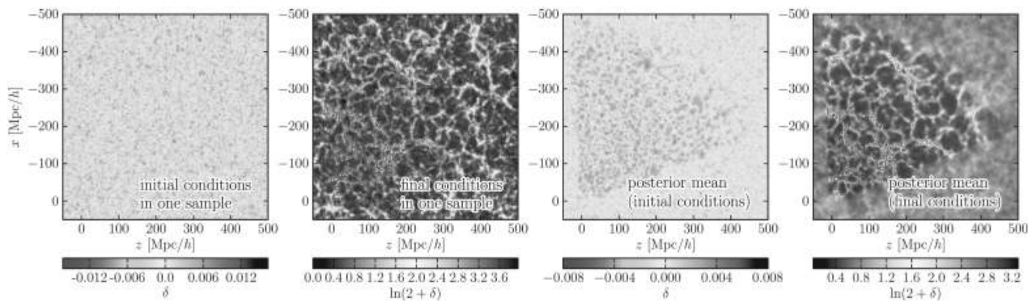

Figure 1. Bayesian LSS inference with BORG in the SDSS DR7. Slices through one sample of the posterior for the initial and final density fields (left) and posterior mean in the initial and final conditions (right). The input galaxies are overplotted on the final conditions as red dots.

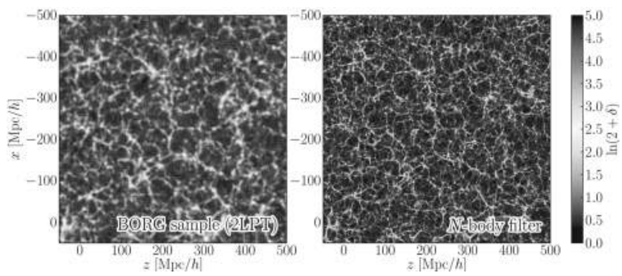

Figure 2. N-body filtering of a BORG sample (left), to produce a non-linear data-constrained realization of the redshift-zero large-scale structure (right).

3.2. *Non-linear filtering*

Building upon these results, it is possible to post-process the samples using fully non-linear dynamics as an additional filtering step (Leclercq *et al.* 2015a). We generate a set of data-constrained realizations of the present large-scale structure: some samples of inferred initial conditions are evolved with 2LPT to $z = 69$, then with a fully non-linear cosmological simulation (using GADGET-2) from $z = 69$ to $z = 0$. This filtering step yields a much more precise view of the deeply non-linear regime of cosmic structure formation, sharpening overdense, virialized structures and resolving more finely the substructure of voids (Fig. 2).

4. Cosmic web analysis

4.1. *Tidal shear classification*

The results presented in § 3.1 form the basis of the analysis of Leclercq *et al.* (2015b), where we classify the cosmic large scale structure into four distinct web-types (voids, sheets, filaments and clusters) and quantify corresponding uncertainties. We follow the dynamic cosmic web classification procedure proposed by Hahn *et al.* (2007), based on the eigenvalues $\lambda_1 < \lambda_2 < \lambda_3$ of the tidal tensor T_{ij}, Hessian of the rescaled gravitational potential: $T_{ij} \equiv \partial^2\Phi/\partial\mathbf{x}_i\,\partial\mathbf{x}_j$, where Φ follows the Poisson equation ($\nabla^2\Phi = \delta$). A voxel is in a cluster (resp. in a filament, in a sheet, in a void) if three (resp. two, one, zero) of the λs are positive.

By applying this classification procedure to all density samples, we are able to estimate the posterior of the four different web-types, conditional on the observations. The means of these pdfs are represented in Fig. 3.

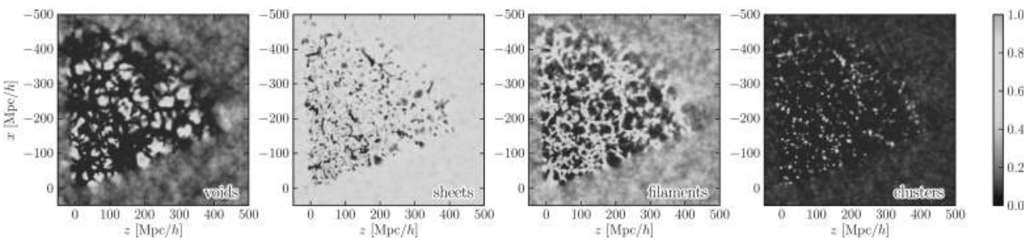

Figure 3. Mean of the posterior pdf for the four different web-types in the SDSS DR7.

4.2. *Dark matter voids*

In Leclercq *et al.* (2015a), we apply computational geometry tools (VIDE: the Void IDentification and Examination pipeline, Sutter *et al.* 2015) to the constrained parts of the non-linear realizations described in § 3.2. We find physical cosmic voids in the field traced by the dark matter particles, probing a level deeper in the mass distribution hierarchy than galaxies. Due to the high density of tracers, we find about an order of magnitude more voids at all scales than the voids directly traced by the SDSS galaxies. In this fashion, we circumvent the issues due to the conjugate and intricate effects of sparsity and biasing on galaxy void catalogs (Sutter *et al.* 2014) and drastically reduce the statistical uncertainty. For usual void statistics (number count, radial density profiles, ellipticities), all the results we obtain are consistent with N-body simulations prepared with the same setup.

Acknowledgements

We thank Jacopo Chevallard, Jens Jasche, Nico Hamaus, Guilhem Lavaux, Emilio Romano-Díaz, Paul Sutter and Alice Pisani for a fruitful collaboration on the projects presented here. FL acknowledges funding from an AMX grant (École polytechnique) and BW from a senior Excellence Chair by the Agence Nationale de la Recherche (ANR-10-CEXC-004-01). This work made in the ILP LABEX (ANR-10-LABX-63) was supported by French state funds managed by the ANR within the Investissements d'Avenir programme (ANR-11-IDEX-0004-02).

References

Adelman-McCarthy, J. K., Agüeros, M. A., *et al.* 2008, *Astrophys. J. Supp.*, 175, 297
Blanton, M. R., Schlegel, D. J., Strauss, M. A., *et al.* 2005, *AJ*, 129, 2562
Bellman, R. E. 1961, *Adaptive Control Processes: A Guided Tour* (Princeton University Press)
Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. 1987, *Physics Letters B*, 195, 216
Hahn, O., Porciani, C., Carollo, C. M., & Dekel, A. 2007, *Mon. Not. R. Astron. Soc.*, 375, 489
Jasche, J. & Wandelt, B. D. 2013, *Mon. Not. R. Astron. Soc.*, 432, 894
Jasche, J. & Wandelt, B. D. 2013, *ApJ*, 779, 15
Jasche, J., Leclercq, F., & Wandelt, B. D. 2015, JCAP, 1, 036
Kitaura, F.-S. 2013, *Mon. Not. R. Astron. Soc.*, 429, L84
Leclercq, F., Jasche, J., Sutter, P. M., Hamaus, N., & Wandelt, B. 2015, *JCAP*, 3, 047
Leclercq, F., Jasche, J., & Wandelt, B. 2015, arXiv:1502.02690
Sutter, P. M., Lavaux, G., Hamaus, N., et al. 2014, *Mon. Not. R. Astron. Soc.*, 442, 462
Sutter, P. M., Lavaux, G., Hamaus, N., et al. 2015, *Astronomy and Computing*, 9, 1
Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., *et al.* 2008, *ApJ*, 674, 1217
Wang, H., Mo, H. J., Yang, X., & van den Bosch, F. C. 2013, *ApJ*, 772, 63

Statistical Challenges in 21st Century Cosmology
Proceedings IAU Symposium No. 306, 2014
A. F. Heavens, J.-L. Starck & A. Krone-Martins, eds.

doi:10.1017/S1743921314013908

Bayesian model comparison in cosmology

Daniel J. Mortlock

Astrophysics Group, Blackett Laboratory, and Department of Mathematics
Imperial College London, London SW7 2AZ, United Kingdom
email: mortlock@ic.ac.uk

Abstract. The standard Bayesian model formalism comparison cannot be applied to most cosmological models as they lack well-motivated parameter priors. However, if the data-set being used is separable, then it is possible to use some of the data to obtain the necessary parameter distributions, the rest of the data being retained for model comparison. While such methods are not fully prescriptive, they provide a route to applying Bayesian model comparison in cosmological situations where it could not otherwise be used.

Keywords. methods: statistical, cosmology: cosmological parameters

1. Introduction

Much of observational cosmology can be thought of as an attempt to use astronomical data to discriminate between the different cosmological models under consideration. Given both the inevitably imperfect data and the intrinsically stochastic nature of many cosmological measurements (*i.e.*, cosmic variance), it is generally impossible to come to absolute conclusions about the various candidate models; the best that can be hoped for is to evaluate the probabilities, conditional on the the available data, that each of the candidate models is the correct description of the Universe. The fact that there is, as far as is known, just a single observable Universe (*i.e.*, there is no ensemble from which it has been drawn), means that such probabilities cannot be frequency-based, and must instead must represent a degree of implication. Self-consistency arguments then require (Cox 1946) that these probabilities be manipulated and inverted using Bayes's theorem.

Taken together, the above facts imply that Bayesian model comparison (Section 2) should be used to assess how well different cosmological models explain the available data, although the fact that most such models have unspecified parameters is a significant difficulty for this approach (Section 3). This problem can be solved for separable data-sets as it is possible to use a two-step method of model comparison (Section 4), illustrated here with high-redshift supernova (SN) data (Section 5).

2. Bayesian model comparison

Given that one of a set of N models, $\{M_1, M_2, \ldots, M_N\}$, is assumed to be true, the state of knowledge conditional on all the available (and relevant) information, I, is fully summarised by the probabilities $\Pr(M_1|I), \Pr(M_2|I), \ldots, \Pr(M_N|I)$, where $\Pr(M_i|I)$ is the probability that the i'th model is correct (and $i \in \{1, 2, \ldots, N\}$). In the light of some new data, d, that has not already been included in the above probabilities, Bayes's theorem gives the updated probability that model i is correct as

$$\Pr(M_i|d, I) = \frac{\Pr(M_i|I) \Pr(d|M_i, I)}{\sum_{j=1}^{N} \Pr(M_j|I) \Pr(d|M_j, I)}, \tag{2.1}$$

where $\Pr(d|M_i, I)$ is the marginal likelihood under model M_i.

Cambridge University Press
978-1-107-07856-7 — Statistical Challenges in 21st Century Cosmology (IAU S306)
Edited by Alan Heavens , Jean-Luc Starck , Alberto Krone-Martins
Excerpt
More Information

If model M_i has N_i unspecified parameters $\{\theta_i\} = \{\theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,N_i}\}$ then the model-averaged likelihood is obtained by marginalising over these parameters to give

$$\Pr(d|M_i, I) = \int \Pr(\{\theta_i\}|M_i, I) \Pr(d|\{\theta_i\}, M_i, I) \, \mathrm{d}\theta_{i,1} \, \mathrm{d}\theta_{i,2} \ldots \mathrm{d}\theta_{i,N_i} \, , \tag{2.2}$$

where $\Pr(\{\theta_i\}|M_i, I)$ is the prior distribution of the parameter values in this model. This expression demonstrates that the full specification of a model requires not just an explicit parameterisation, but a distribution for those parameters as well; two mathematically identical descriptions with different parameter priors are, in fact, different models.

3. Comparison of models without parameter priors

Equations 2.1 and 2.2 together summarise a self-consistent method for assessing which of a set of models is better supported by the available information, provided that the parameter priors for all the models are explicitly defined and unit-normalised. In particular, while it is often possible to obtain sensible parameter constraints based on an improper prior, such as $\Pr(\{\theta_i\}|M_i, I)$ constant for all $\{\theta_i\}$, the resultant marginal likelihood is meaningless (Dickey 1961). Unfortunately, it is commonly the case in astronomy and cosmology that there is no compelling form for the models' parameter priors and, further, that the natural uninformative prior distributions are improper and cannot be normalised. The apparent implication is that Bayesian model comparison, at least in the form described in Section 2, cannot be used in cosmology, an idea that has been explored previously by, *e.g.*, Efstathiou (2008) and Jenkins & Peacock (2011). The disturbing corollary would be that there is no self-consistent method to choose between the available cosmological models, even if they are completely quantitative and mathematically well-defined.

4. Model comparison with separable data

The idea that the relative degree of support for models with unspecified parameters is undefined is at odds with the marked – and data-driven – progress that has been made in cosmology over the last century. Clearly it *is* possible to use data to choose sensibly between models even if they do not have well-motivated parameter priors; but can this be formalised in a way that satisfies Bayes's theorem and is hence logically self-consistent?

One possibility is, for separable data-sets (such as those which consist of measurements of many astronomical sources), to use some of the available data to obtain the necessary parameter priors and to then use the remaining data for model comparison. This is an old concept, dating back at least to Lempers (1971) and explored subsequently by, *e.g.*, Spiegelhalter & Smith (1982) and O'Hagan (1995). The central idea is to partition the data as $d = (d_1, d_2)$, with the first partition of training data used to obtain the (partial) posterior distribution for the parameters of i'th model as

$$\Pr(\{\theta_i\}|d_1, M_i, I) = \frac{\Pr(\{\theta_i\}|M_i, I) \Pr(d_1|\{\theta_i\}, M_i, I)}{\int \Pr(\{\theta'_i\}|M_i, I) \Pr(d_1|\{\theta'_i\}, M_i, I) \, \mathrm{d}\theta'_{i,1} \, \mathrm{d}\theta'_{i,2} \ldots \mathrm{d}\theta'_{i,N_i}}, \tag{4.1}$$

where $\Pr(\{\theta_i\}|M_i, I)$, which need *not* be normaliseable, should be a highly uninformative prior. This posterior distribution can then be used as the prior needed to obtain a meaningful marginal likelihood, which can then be evaluated for the testing data as

$$\Pr(d_2|d_1, M_i, I) = \int \Pr(\{\theta_i\}|d_1, M_i, I) \Pr(d_2|\{\theta_i\}, M_i, I) \, \mathrm{d}\theta_{i,1} \, \mathrm{d}\theta_{i,2} \ldots \mathrm{d}\theta_{i,N_i} \, . \tag{4.2}$$

Cambridge University Press
978-1-107-07856-7 — Statistical Challenges in 21st Century Cosmology (IAU S306)
Edited by Alan Heavens , Jean-Luc Starck , Alberto Krone-Martins
Excerpt
More Information

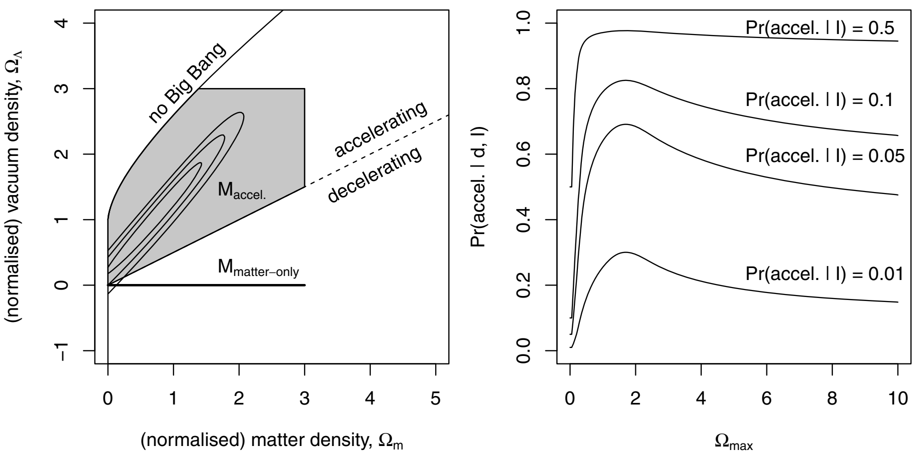

Figure 1. (left) The posterior distribution of Ω_m and Ω_Λ implied by from the Perlmutter *et al.* (1999) SCP SN data and a uniform prior with $\Omega_\mathrm{m} \geqslant 0$. Highest posterior density contours enclosing 68.3%, 95.4% and 99.7% of the posterior probability are shown. Also shown are the prior distributions of the accelerating model and matter only model for $\Omega_\mathrm{max} = 3$. (right) The dependence of $\Pr(\mathrm{accel.}|d, I)$ on Ω_max, shown for different prior probabilities, $\Pr(\mathrm{accel.}|I)$.

This marginal likelihood is coherent, in the sense that it provides self-consistent updated posterior probabilities when inserted into Equation 2.1, but there is also ambiguity: there is no compelling scheme for partitioning the data. It is tempting to average over the possible partitions, but this approach does not have a rigorous motivation. Despite these ambiguities, this two-step method of Bayesian model comparison for separable data does satisfy the Cox (1946) self-consistency requirements and so provide a means of calculating posterior probabilities for cosmological models with unspecified parameter priors.

5. Example: late-time acceleration and supernovae

One of the most significant recent cosmological discoveries was that the Universe's expansion rate is increasing, a result which is often linked most strongly to the observations of distant SNe made by Riess *et al.* (1998) and Perlmutter *et al.* (1999). The comparative faintness of the SNe, given their redshifts and light-curve decay timescales, indicated that the (normalised) cosmological constant, Ω_Λ, is sufficiently large to override the deceleration caused by the (normalised) matter density, Ω_m. Riess *et al.* (1998) and Perlmutter *et al.* (1999) used their SNe measurements, d, to obtain posterior distributions of the form $\Pr(\Omega_\Lambda, \Omega_\mathrm{m}|d, I)$, under the assumption of unimformative (and improper) uniform priors of the form $\Pr(\Omega_\mathrm{m}, \Omega_\Lambda) \propto \Theta(\Omega_\mathrm{m})$, where $\Theta(x)$ is the Heaviside step function. The posterior distribution for the 42 SCP SNe from Perlmutter *et al.* (1999), reproduced in Fig. 1, reveals that most of the models that are consistent with the data correspond to an accelerating universe (*i.e.*, $\Omega_\Lambda > \Omega_\mathrm{m}/2$).

But do these data provide *quantitive* evidence of cosmological acceleration? Riess *et al.* (1998) approached this question by calculating the fraction of the posterior with $\Omega_\Lambda > \Omega_\mathrm{m}/2$, which is an apparently compelling 0.997 for the case shown in Fig. 1. The relevant Bayesian calculation (*c.f.* Drell *et al.* 2000) should, however, be based on the marginal likelihoods of an accelerating model (for which the prior is non-zero only for $\Omega_\Lambda > \Omega_\mathrm{m}/2$) and a decelerating model (for which the obvious option is a matter-only model with $\Omega_\Lambda = 0$). Such models can be fully specified (in the sense defined in Section 2) by adding the restrictions that $0 \leqslant \Omega_\mathrm{m} \leqslant \Omega_\mathrm{max}$ and $0 \leqslant \Omega_\Lambda \leqslant \min[\Omega_\mathrm{max}, \Omega_{\Lambda,\mathrm{BB}}(\Omega_\mathrm{m})]$ (defined to

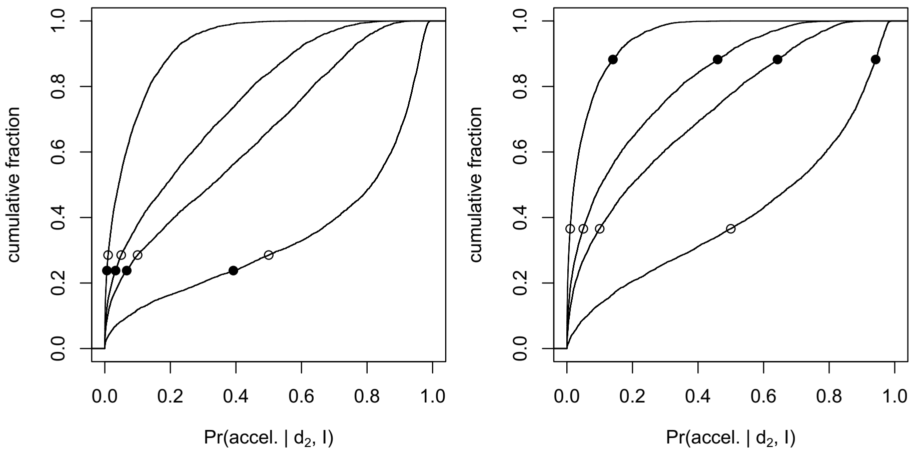

Figure 2. The distribution of $\Pr(\text{accel.}|d_2, I)$ obtained from different partitions of the Perlmutter *et al.* (1999) SN data set with training sets of 10 (left) and 21 (right) SNe. The open symbols indicate the prior values (of, from left to right, 0.01, 0.05, 0.1 and 0.5) and the solid symbols show the posterior values given by training and testing samples that alternate in redshift.

reject models that did not begin with a Big Bang), where $\Omega_{\rm max} \geqslant 0$ is an unspecified "hyper-parameter". Figure 1 shows the dependence of the posterior probability of the accelerating model, $\Pr(\text{accel.}|d, I)$, on $\Omega_{\rm max}$. Even the peak values of $\Pr(\text{accel.}|d, I)$ are considerably lower than the posterior fraction quoted above, and the dependence on the unknown value of $\Omega_{\rm max}$ is significant as well.

Rather than introducing an arbitrary new parameter, another option is to adopt the two-step method described in Section 4, using some of the SN data to obtain a partial posterior in $\Omega_{\rm m}$ and Ω_Λ for both the accelerating and matter-only models and then using the remainder to perform model comparison. The results of doing so are shown in Fig. 2 for several different partitioning options. These results again illustrate the standard Bayesian result that the better-fitting accelerating model is not favoured so decisively over the more predictive (*i.e.*, "simpler") matter-only model, a result that robust to prior choice.

This two-step approach to model comparison could be applied to a variety of problems in astrophysics and cosmology (*e.g.*, Bailer-Jones 2012, Khanin & Mortlock 2014).

References

Bailer-Jones, C. 2012 *Astronomy & Astrophysics*, 546, A89
Cox, R. T. 1946, *American Journal of Physics*, 14, 1
Drell, P. S., Loredo, T. J., & Wasserman, I. 2000, *The Astrophysical Journal*, 530, 593
Efstathiou, G. 2008, *Monthly Notices of the Royal Astronomical Society*, 388, 1314
Jenkins, C. R. & Peacock, J. A. 2011, *Monthly Notices of the Royal Astronomical Society*, 413, 2895
Khanin, A. & Mortlock, D. J. 2014, *Monthly Notices of the Royal Astronomical Society*, 444, 1591
Lempers, F. B. 1971, *Posterior Probabilities of Alternative Linear Models*, Rotterdam: University Press
O'Hagan, A. 1995, *Journal of the Royal Statistical Society B*, 30, 490
Perlmutter, S., *et al.* 1999, *The Astrophysical Journal*, 517, 565
Spiegelhalter, D. J. & Smith, A. F. M., 1982, *Journal of the Royal Statistical Society B*, 44, 377
Riess, A. G., *et al.* 1998, *The Astronomical Journal*, 116, 1009

Statistical Challenges in 21st Century Cosmology
Proceedings IAU Symposium No. 306, 2014
A. F. Heavens, J.-L. Starck & A. Krone-Martins, eds.

doi:10.1017/S1743921314010722

What we talk about when we talk about fields

Ewan Cameron

Department of Zoology, University of Oxford, Tinbergen Building, South Parks Road, Oxford, OX1 3PS, United Kingdom
email: dr.ewan.cameron@gmail.com
website: astrostatistics.wordpress.com

Abstract. In astronomical and cosmological studies one often wishes to infer some properties of an infinite-dimensional field indexed within a finite-dimensional metric space given only a finite collection of noisy observational data. Bayesian inference offers an increasingly-popular strategy to overcome the inherent ill-posedness of this signal reconstruction challenge. However, there remains a great deal of confusion within the astronomical community regarding the appropriate mathematical devices for framing such analyses and the diversity of available computational procedures for recovering posterior functionals. In this brief research note I will attempt to clarify both these issues from an "applied statistics" perpective, with insights garnered from my post-astronomy experiences as a computational Bayesian / epidemiological geostatistician.

Keywords. Methods: data analysis–methods: statistical

1. Introduction

The potential afforded by Bayesian techniques for inferring the properties of infinite-dimensional mathematical structures, such as random fields (to be understood here as random functions defined at each point of some finite-dimensional metric space), has long been recognised by both probability theorists, e.g. O'Hagan (1978), and practitioners: with the first wave of practical applications in geoscience (e.g. Omre (1987), Handcock & Stein (1993)) and machine learning (e.g. Rasmussen & Williams (1996), Neal (1997)) contemporaneous with the advent of sufficiently powerful desktop computers. Cosmologists were at this time notable as *'early adopters'* and pioneers of the new techniques for field inference. Indeed, the Monte Carlo methods for constrained simulation from Gaussian random fields developed by Bertschinger (1987) and Hoffman & Ribak (1991) remain key tools for efficient conditional simulation, cf. Doucet (2010).

However, over the past decade the sophistication of statistical analysis techniques brought to bear on the study of cosmological fields has not kept pace with progress outside of astronomy. With modern tools such as the Integrated Nested Laplace Approximation (INLA; Rue *et al.* (2009)), 'variational inference' (Hensman *et al.* (2013)), particle filtering (Del Moral *et al.* (2007)), and Approximate Bayesian Computation (ABC; Marjoram *et al.* (2003)) almost entirely ignored to-date by the cosmological community we have, in my opinion, become the *'laggards'* of the technology adoption lifecycle.

There are multiple factors seemingly to blame for this divergence: (i) the emergence of an isolationist attitude to the practice of cosmological statistics; (ii) an over-emphasis on the path integration-based conceptulisation of random fields, rather than the measure-theory-based mathematics of mainstream statistics; and (iii) an under-appreciation of the potential for stochastic process priors (including, *but not limited to*, the Gaussian process) as flexible modelling components within the hierarchical Bayesian framework. With the first already being well fought back against by inter-disciplinary programming in conferences such as the IAUS306 and the SCMA series I will therefore focus in this proceedings (as in my contributed talk) on the latter two. In particular, I aim to clarify a number of mathematical concepts crucial to a high-level understanding of Bayesian inference over random fields and measures (Section 2), and then to highlight just a few of the exciting techniques to have recently emerged in this area (Section 3).

2. The Mathematics of Bayesian Field Inference

Cosmologists and astronomers already well-versed in the practice and theory of Bayesian statistics in the finite-dimensional setting will typically have one of two contrasting experiences upon first attempting to extend these ideas to infinite-dimensional inference problems. The pragmatist will happily observe that the mechanics of computation are little changed (e.g. the Gaussian random field at finite sample points is distributed just as the familiar multivariate Gaussian), while the cautious theorist will more likely be overwhelmed by a first acquaintance with measure-theoretic probability (i.e., probability triples and the algebra of sets). But ideally one will have both experiences, since each offers an equally important perspective, as I discuss in this Section.

2.1. *Distributions over Infinite-Dimensional Space*

In formal statistics the core of probabilistic computation is framed within the language of measure theory: the key object being the *'probability triple'* of (i) a sample space, Ω, i.e., some non-empty set; (ii) a σ-algebra, Σ, i.e., a collection of subsets of Ω with $\emptyset, \Omega \in \Sigma$, closed under the formation of complements, countable unions and intersections; and (iii) a probability measure, P, i.e., a countably additive set function from Σ to [0,1] for which $P(\emptyset) = 0$ and $P(\Omega) = 1$. In this context Carathéodory's Extension Theorem provides the theorist with the machinary to build complex probability triples and forge a rigorous notion of random variables as measures on the pre-images of Borel σ-algebra sets of the real numbers; and from this to the familiar mechanics of probability densities defined with respect to the Lebesgue measure (e.g. the standard Normal with $f(x) = \frac{1}{\sqrt{2\pi}} \exp -\frac{x^2}{2} dx$). Nevertheless, with the Lebesgue measure behaving intuitively as a product measure in $\mathcal{R}^n$, and with Lebesgue and Riemann integration interchangable in practice for all but a few rare cases, the pragmatist can safely ignore these theoretical foundations in the study of 'real-world' problems in finite-dimensional settings.

In the context of probabilistic inference over *fields*, however, one must proceed with care as there exists no equivalent to the Lebesgue measure to serve as a natural reference for defining densities in an infinite-dimensional Banach space (e.g. the L^p function spaces). Hence, for Bayesian analysis in infinite-dimensional space we must be deliberate in our choice of reference measure, which we encode into the prior. Typically we will do this indirectly by assigning as prior the implicit measure (or 'law') belonging to a given *stochastic process* (e.g. a Gaussian process, or Poisson process) having sample paths within the field space under study. Although quite technical the distinction between this formal statistical approach and the path integration-based language of cosmological papers, e.g. Enßlin *et al.* (2009), Kitching & Taylor (2011), is important if we are to connect with, *and thereby benefit from*, the rich body of applied statistics literature on infinite-dimensional inference. Worth noting also is that the measure-theoretic equivalent of the probability density is the 'Radon-Nikodym (R-N) derivative', with a trivial but illustrative example being that of the R-N derivative of posterior against prior given by the likelihood function divided by the marginal likelihood, c.f. Cotter *et al.* (2009).

Finally, the measure theoretic definition of a stochastic process is a collection of random variables *indexed* by a set; here all points of the physical metric space over which