



HAL
open science

A special issue on Bayesian Inference: Challenges, Perspective, and Prospects

Christian Robert, Judith Rousseau

► **To cite this version:**

Christian Robert, Judith Rousseau. A special issue on Bayesian Inference: Challenges, Perspective, and Prospects. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2023, 381 (2247), <10.1098/rsta.2022.0155>. <hal-04049902>

HAL Id: hal-04049902

<https://hal.science/hal-04049902v1>

Submitted on 29 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

A special issue on Bayesian Inference: Challenges, Perspective, and Prospects

Christian P. Robert¹ and Judith Rousseau²
University of ^{1,2}Paris Dauphine, ²Oxford, and ¹Warwick

February 22, 2023

This special issue is dedicated to Sir Adrian Smith, whose contributions to Bayesian analysis have deeply impacted the field (or rather fields of Bayesian inference, decision theory, and statistical computing). When contemplating his past achievements, it is striking to align the emergence of massive advances in these fields with some papers or books of his. For instance, Lindley’s and Smith’s “Bayes Estimates for the Linear Model” (1971), a Read Paper at the Royal Statistical Society, is making the case for the Bayesian analysis of the most standard statistical model, as well as emphasizing the notion of exchangeability that is foundational in Bayesian statistics, paving the way to the emergence of hierarchical Bayesian modelling. It thus makes the link between the early days of Bruno de Finetti, whose work Adrian Smith translated into English and the current research in non-parametric and robust statistics. Bernardo’s and Smith’s masterpiece, *Bayesian Theory* (1994), sets statistical inference within decision- and information-theoretic frameworks in a most elegant and universal manner that could be deemed a Bourbaki volume for Bayesian statistics if this endeavour had reached further than pure mathematics. It also emphasizes the central role of hierarchical modelling in the construction of priors, as exemplified in Carlin’s, Gelfand’s, and Smith’s “Hierarchical Bayesian analysis of changepoint problems” (1992).

The series of papers published in 1990 by Alan Gelfand and Adrian Smith, esp. “Sampling-Based Approaches to Calculating Marginal Densities” (1990), is overwhelmingly perceived as the birth date of modern Markov chain Monte Carlo (MCMC) methods, as it brought to the whole statistics community (and then quickly wider communities) the realisation that MCMC simulation was the sesame to unlock complex modelling issues. The consequences on the adoption of Bayesian modelling by non-specialists are enormous and long-lasting.

Similarly, Gordon’s, Salmond’s, and Smith’s, “Novel approach to nonlinear/non-Gaussian Bayesian state estimation” (1992) is considered as the birthplace of sequential Monte Carlo, aka particle filtering, with considerable consequences in tracking, robotics, econometrics, and many other fields.

Titterton’s, Smith’s, and Makov’s reference book, “Statistical Analysis of Finite Mixtures” (1984) is a precursor in the formalisation of heterogeneous

data structures, paving the way for the incoming MCMC resolutions like Tanner and Wong (1987), Gelman and King (1990), and Diebolt and Robert (1990). Further, and many other fields. Denison's, Holmes's, Mallick's, and Smith's book, "Bayesian methods for nonlinear classification and regression" (2002) is another testimony to the influence of Adrian Smith on the field, stressing the emergence of robust and general classification and non-linear regression methods to analyse complex data, prefiguring in a way the later emergence of machine learning methods, with the additional Bayesian assessment of uncertainty. It is also bringing forward the ability of Bayesian non-parametric modelling that is now broadly accepted, following a series of papers by Denison, Mallick and Smith in the late 1990's like CART and MARS.

Peter Grünwald proposes a decision-based construction of Bayesian inference and uncertainty quantification via e-posteriors that bypasses the precise construction of a prior distribution. It is advanced as a more robust form of inference, with guaranteed frequentist properties.

Chris Holmes and Stephen Walker connect exchangeability with Bayesian bootstrap. Exchangeability is at the core of Bayesian thinking, as introduced by de Finetti (1937) and emphasized in Bernardo and Smith (2009). Their perspective reinforces this positioning by explaining exchangeability through learning, which is the main argument for pursuing the Bayesian approach to statistics, while connecting with a more classical approach by predicting an infinite population from the current dataset. By extending the notion of Bayesian bootstrap, they reformulate the Bayesian paradigm as purely predictive.

Sonia Petrone also focus on the predictive aspect of Bayesian inference, making the crucial point that it is a natural way to connect with machine learning imperatives. She makes the method naturally embedded within nonparametric statistics, as a form of Pólya urn sampling, and establishes limit coverage properties that grant frequentist properties to the ensuing Bayesian procedures.

Veronika Rockova and Lizhen Nie brings another perspective on the Bayesian bootstrap. The approach there is more likelihood-free in the sense that the sampling model is no longer defined through a likelihood but instead through a loss function as in Bissiri et al. (2016). A strong innovation in this approach is to rely on machine-learning based tool, a generative sampler that allows for a generalisation of the Bayesian bootstrap of Newton and Raftery (1994) that brings both computational advantages and theoretical guarantees, including under model mis-specification, a perspective under-studied in the field.

When seeking computational solutions to approximate an essential Bayesian entity, namely the Bayes factor (or the evidence), Arnaud Doucet, Eric Mouline and Achille Thin similarly exploit machine-learning tools like variational auto-encoders and normalizing flows to build rich families of approximations in settings where the likelihood function is unavailable. They bring in an annealing perspective to improve further the computational efficiency of these importance sampling methods.

Alicia Carriquiry and Kori Khan build a safer environment for forensic calibration studies by resorting to hierarchical Bayes modelling, investigated by AFMS in his early career, and borrowing from small area estimation, thus

bringing a quantitative evaluation of the imprecision in using this tool for reporting error rates.

Matthew Stephens brings a broader perspective to the volume by connecting Bayesian principles with non-Bayesian procedures, like confidence intervals and p-values, also connecting with Peter Grünwald e-values. He also advocates for considering alternative approaches for faster or more robust resolutions.

Sylvia Frühwirth-Schnatter broadens the range of shrinkage priors in connection with spike-and-slab priors used in variable selection and factor analysis. She reveals new connections with Bayesian non-parametrics and exchangeability. The computational aspects are also most relevant for the theme of this volume since they involve MCMC and data augmentation, where AFMS's contributions have been fundamental.

Sara Wade advocates for the Bayesian approach to cluster analysis, of which AFMS has been an early researcher when working on mixtures of distributions (Titterington, Smith and Makov, 1984). She brings another connection with the Gibbs priors of Bissiri et al. (2016).

Kerrie Mengersen and her coauthors also offers a very broad perspective on the opportunities and challenges of applied Bayesian analysis, including design (with links to decision theory and hence AFMS), data sources and citizen science, federated learning and data privacy, with MCMC consequences, implicit models and ABC, and model transfer. All these points are illustrated by diverse and exciting case studies.

Richard Nickl and coauthors study hindrances to MCMC convergence in high dimensions, reconnecting with the early days when MCMC and simulated annealing were created in parallel and phase transition sounded like an epiphenomenon. The notion of entropy wells is quite appealing in demonstrating that convergence issues can arise with MCMC even in the case of unimodal (if not log-concave) targets. Giving further arguments against local Monte Carlo methods and possibly in favour of new schemes like PDMPs although none is included in this paper. Too bad notations are not always defined precisely, like $\{a, b\}$. For a specific Gaussian tensor model, the posterior concentrates exponentially fast (with the dimension) over a small region of the space, inducing an equally exponential time for "small step" MCMC algorithms to converge. This includes both unadjusted and Metropolis-adjusted Langevin algorithms.

Peter Müller, Noirit Chandra, and Abhra Sarkar open a new perspective on Bayesian exploitation of complex data sources like medical record. They demonstrate how Bayesian non-parametrics can be used in that context, with consequence on drug trial. The approach relies on MCMC methods, of which AFMS was a founding father. The paper also relates to his input on mixtures and Bayesian non-parametrics.

Sumio Watanabe presents his original perspective on Bayesian information criteria (BIC), within the context of cross-validation, including new results on the asymptotic behaviour of Bayes factors, new types of losses and BICs. The approach defended therein also allows for an assessment of prior distributions and a novel form of empirical Bayes analysis by optimising hyperparameters. Potential link with ELBOs of Doucet et al.

Beatrice Franzolini and her coauthors study an integer-supported zero-inflated model called the hurdle model, with the addition of a clustering mechanism, connecting with Wade's article. The work builds on earlier advances on finite mixtures with random number of components, which brings a higher computational efficiency in the Bayesian analysis of the model than when using non-parametric alternatives.