Sequential Neural Likelihood and Posterior Approximation: Inference for Intractable Probabilistic Models via Direct Density Estimation

Samuel Wiqvist[1], Jes Frellsen[2] and Umberto Picchini[3]

[1] *Centre for Mathematical Sciences, Lund University, Lund, Sweden, samuel.wiqvist@matstat.lu.se*

[2] *Department of Applied Mathematics and Computer Science, DTU, Denmark, jefr@dtu.dk*

[3] *Department of Mathematical Sciences at Chalmers University of Technology and University of Gothenburg, Sweden, picchini@chalmers.se*

We introduce the *sequential neural posterior and likelihood approximation* (SNPLA) algorithm [1]. SNPLA produces Bayesian inference in implicit models and utilizes direct density estimation schemes to simultaneously learn both the posterior and the likelihood function of the model. An implicit model is one such that the likelihood is unknown in closed-form, but can be learned via computer-assisted simulations. Therefore SNPLA is an example of simulation-based inference methodology, a field that has gained much traction in recent years. Specifically, our work uses deep learning architectures, such as normalizing flows. We will discuss how SNPLA connects with some recent developments in the field of *machine-learning-based* inference for implicit models. Over several examples, we show that SNPLA obtains inference results that are similar to its competitors, while utilizing the same computational resources, even though the inference problem for SNPLA is particularly complex, since its goal is not only to produce posterior draws for the parameters of interest, but simultaneously to learn an approximation to the likelihood function.

References

[1] Wiqvist S, Frellsen J., Picchini U. (2021). Sequential Neural Posterior and Likelihood Approximation, *arXiv preprint*, arXiv:2102.06522.

Parameter elimination in particle Gibbs sampling

Anna Wigren[1]

[1] *Uppsala University, Sweden, anna.wigren@it.uu.se*

Bayesian joint parameter and state inference in non-linear state-space models is a difficult problem due to the often high-dimensional state sequence. Particle Gibbs (PG) and its extension particle Gibbs with ancestor sampling (PGAS), both in the particle MCMC family of methods [1], are well-suited for solving this type of inference problem. However, due to their construction, all Gibbs-type methods produce correlated samples. We suggest to reduce the correlation by marginalizing out one or more parameters from the state update. This results in a non-Markovian model, which is intractable in the general case. We show that for models with conjugacy relations between the parameter prior and the complete data likelihood, it is possible to derive marginalized versions of PG and PGAS that scale linearly with the number of observations [2].

It turns out that deriving the marginalized conjugacy relations is quite cumbersome, but we can employ probabilistic programming to do automatic marginalization. The methods presented can provide a performance improvement also when only some of the parameters have a conjugacy relationship, something that greatly increases the applicability in practice.

We will also discuss how the marginalization framework presented above can be extended to hierarchical models. Consider the case when there are several datasets available from processes with the same model structure, but where only a subset of the parameters are shared between the datasets. One example is the spread of a disease in different locations, where parameters like the incubation time are shared, but parameters like the rate of infection may differ across locations. Joint identification for these models with shared parameters can be beneficial, in particular if the individual datasets are small.

References

[1] Andrieu, C., Doucet, A., Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **72(3)**, 269 – 342.

[2] Wigren, A., Risuleo, R. S., Murray, L. M., and Lindsten, F (2019). Parameter elimination in particle Gibbs sampling. *Advances in Neural Information Processing Systems 32.*

Neural Ratio Estimation for Simulation-Based Inference

Gilles Louppe[1]

[1] *University of Liège, Belgium, g.louppe@uliege.be*

Posterior inference with an intractable likelihood is becoming an increasingly common task in scientific domains which rely on sophisticated computer simulations. Typically, these forward models do not admit tractable densities forcing practitioners to make use of approximations. This work introduces a novel approach to address the intractability of the likelihood and the marginal model. We achieve this by learning a flexible amortized estimator which approximates the likelihood-to-evidence ratio. We demonstrate that the learned ratio estimator can be embedded in MCMC samplers to approximate likelihood-ratios between consecutive states in the Markov chain, allowing us to draw samples from the intractable posterior. Techniques are presented to improve the numerical stability and to measure the quality of an approximation. The accuracy of our approach is demonstrated on a variety of benchmarks against well-established techniques. Scientific applications in physics show its applicability.

Spectral density-based and measure-preserving ABC for partially observed diffusion processes.

Evelyn Buckwar[1], **Massimiliano Tamborrino**[2] **and Irene Tubikanec**[3]

[1] *Johannes Kepler University Linz, Austria, evelyn.buckwar@jku.at*
[2] *University of Warwick, England, massimiliano.tamborrino@warwick.ac.uk*
[3] *Johannes Kepler University Linz, Austria, irene.tubikanec@jku.at*

Approximate Bayesian computation (ABC) has become one of the major tools of likelihood-free statistical inference in complex mathematical models. Simultaneously, stochastic differential equations (SDEs) have developed to an established tool for modelling time-dependent, real-world phenomena with underlying random effects. When applying ABC to stochastic models, two major difficulties arise. First, the derivation of effective summary statistics and proper distances is particularly challenging, since simulations from the stochastic process under the same parameter configuration result in different trajectories. Second, exact simulation schemes to generate trajectories from the stochastic model are rarely available, requiring the derivation of suitable numerical methods for the synthetic data generation. To obtain summaries that are less sensitive to the intrinsic stochasticity of the model, we propose to build up the statistical method (e.g. the choice of the summary statistics) on the underlying structural properties of the model. Here, we focus on the existence of an invariant measure and we map the data to their estimated invariant density and invariant spectral density. Then, to ensure that these model properties are kept in the synthetic data generation, we adopt measure-preserving numerical splitting schemes. The derived property-based and measure-preserving ABC method is illustrated on the broad class of partially observed Hamiltonian type SDEs, both with simulated data and with real electroencephalography data. The derived summaries are particularly robust to the model simulation, and this fact, combined with the proposed reliable numerical scheme, yields accurate ABC inference. In contrast, the inference returned using standard numerical methods (Euler-–Maruyama discretisation) fails. The proposed ingredients can be incorporated into any type of ABC algorithm and directly applied to all SDEs that are characterised by an invariant distribution and for which a measure-preserving numerical method can be derived.

References

[1] Buckwar, E., Tamborrino, M. and Tubikanec, I. (2020). Spectral density-based and measure-preserving ABC for partially observed diffusion processes. An illustration on Hamiltonian SDEs, Statistics and Computing, **30**, 627 -– 648. `https://doi.org/10.1007/s11222-019-09909-6`

Guided sequential schemes for intractable Bayesian models

Umberto Picchini[1]

[1] *Chalmers University of Technology and University of Gothenburg, Sweden, picchini@chalmers.se*

Sequential sampling schemes such as sequential importance sampling (SIS) and especially sequential Monte Carlo (SMC) have proved fundamental for Bayesian inference in models not admitting a readily available likelihood function. For approximate Bayesian computation (ABC) sequential Monte Carlo ABC is the state-of-art sampler, however it usually comes at a non-negligible computational cost due to the accept-reject steps applied to possibly many samples from the generative model (particles). We contribute to the ABC modeller's toolbox by providing proposal samplers that are conditional to summary statistics of the data. In a sense, the proposed particles are "guided" to reach regions of the posterior surface that is compatible with observed data. This improves the acceptance rates of these sequential samplers, thus reducing the computational effort, and we studied how to preserve the accuracy in the inference. We provide guided samplers for both SIS-ABC and SMC-ABC. In particular, our guided samplers are competitive in a case-study involving cell movements with high-dimensional summary statistics (dimension 145 to over 400).