Low signal to noise state space modelling using simulation based inference

Ingo P. Waldmann, Marco Rocchetto, Marcel Debczynski & Josh Ferguson
Spaceflux Ltd

ABSTRACT

The challenges of detecting and characterizing very small objects in Medium Earth Orbit (MEO), Geostationary Orbit (GEO), or beyond GEO (xGEO) regimes are significant and have been recognized as an area of strategic importance. Developing high-precision Space Situational Awareness (SSA) solutions is critical in addressing these challenges, especially as current methodologies often fall short in certain key aspects. Optical systems, while adaptable to distant orbit regimes, face limitations due to low signal-to-noise ratios. Not enough light reflected off the Resident Space Object (RSO) reaches the telescope to be detectable in an exposure or frame. Various track-before-detect (TBD) approaches have been proposed in the literature, largely based on variations of hidden Markov models. Approaches such as Kalman filters and multi-Bernoulli filters, as well as variations thereof, provide excellent solutions for point estimation of the maximum a posteriori solutions of the underlying state space model. Another approach to state space modeling is through posterior sampling methods such as Markov Chain Monte Carlo (MCMC). While significantly more burdensome computationally, sampling techniques provide a richer understanding of the state space model's posterior distribution. This can be particularly useful in the presence of non-Gaussian noise and non-linear state space models. This paper focuses on a recent development in sampling theory, simulation-based inference (SBI). We show that SBI techniques, in particular neural posterior estimation, can not only accurately constrain the underlying orbital state space model in very low signal-to-noise conditions but can also be shown to be computationally less intensive than classic MCMC approaches. This allows for a full recovery of the state space's posterior distribution while making the approach sufficiently computationally efficient to be used for regular post-processing of state space models.

1. INTRODUCTION

The tracking of resident space objects (RSOs) spans a very wide range of orbital regimes, ranging from atmosphere-grazing very low Earth Orbits (vLEOs) to orbits spanning the Earth-Moon distance, so-called cislunar. Whilst the accurate tracking of RSOs poses significant challenges across all orbital regimes, we here focus on the identification of very low signal-to-noise measurements of cis-lunar orbits in optical measurements. In low SNR regimes, the individual RSO may not be observable in an individual frame but stacking of frames is required to build up sufficient SNR to detect the RSO (e.g. [17]). Though common practice, this approach is suboptimal when the exact orbital trajectory is not known a priori or when observational noise is significantly non-Gaussian. An alternative approach is to opt for track-before-detect (TBD) methods whereby the data is not coadded but an orbital model is minimised over a time series of noisy frames to reveal the underlying low-SNR RSO and its orbital trajectory (e.g. [12, 5, 15, 11, 4]). Much excellent work has recently been published on using various flavours of Kalman Filters or Random Finite Set Statistics (e.g. [17, 10, 13, 2]) such as multi-Bernoulli Filters (e.g. [19, 18, 7, 14, 8]).

Here, we explore an alternative approach to these techniques and attempt to model the full Bayesian posterior distribution of the state space model (i.e., the a priori unknown orbital solution of the RSO) using a simulation-based inference approach. Typically, deriving a full posterior distribution requires a sampling-based approach when the likelihood is not analytically tractable. Deriving the full posterior distribution of the orbital state space model can be particularly useful in the presence of non-Gaussian observational noise. Here, a full analysis of the posterior can provide valuable insights into the robustness of the orbital solution and state space model degeneracies resulting from weakly constraining low SNR measurements. This is particularly advantageous for so-called 'stare' campaigns where the original orbital solution of the RSO is poorly or unconstrained. Here, a full exploration of the orbital state-vector parameter space is required. The problem with sampling-based approaches such as Markov Chain Monte Carlo (MCMC, [1]) or Nested Sampling ([16, 6]) is the often prohibitive computational expense of sampling the posterior distribution. Hence, these approaches have not been widely studied in the context of orbital determination, where a large data rate requires near real-time performance.

In this paper, we showcase a proof of concept (PoC) study of using amortized sequential neural posterior estimation (SNPE,[3]) to alleviate most of the previously mentioned sampling concerns and to achieve real-time performances in estimating posterior distributions in low SNR observations.

2. SIMULATION BASED INFERENCE

Here we will briefly introduce the concept of Neural Posterior Estimators and how they can be used to provide accurate estimates of posterior distributions without the need to sample the posterior each time a new data set is provided.

We define the standard Bayesian argument as

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} \tag{1}$$

where $p(\theta|x)$ is the posterior distribution of the orbital state space model with parameters θ given the observed data x. For simplicity of notation, we do not differentiate between vectors of parameters and individual model parameters here; however, in most cases, these are vectors. Similarly, $p(x|\theta)$, $p(\theta)$ and $p(x)$ are the likelihood, prior distribution and Bayesian Evidence, respectively. Instead of sampling from likelihood and prior directly, such as is the case with MCMC techniques, simulation-based-inference (SBI) foregoes this sampling step and trains a neural network to map between prior drawn samples of parameters, θ, and corresponding simulated data samples, x. This assumes that the orbital state can be simulated by a forward model. By learning the mapping between θ and x we implicitly learn the underlying shape of the likelihood function and hence the posterior distribution. As this mapping can be learned for all parameters within the prior volume, we are able to learn an approximation of the posterior distribution for all instances of θ. In other words, we amortize our Bayesian inference over the full prior volume and we do not need to sample the posterior distribution anew for a new data set[1]. Since this is a likelihood free method, we do not need to specify the degree or nature of non-Gaussian noise in our data. Another advantage of the SBI approach.

SBI is a rapidly growing field of research and we refer the reader to ([3] and [20]) for a more in-depth discussion. Here, we follow the formalism by [9].

The mapping of $\theta - > x$ using a neural network can be described as conditional density estimation. For each data point x_0 we compute an approximation of the posterior distribution through having learned all possible mappings between x and θ. We denote this approximate posterior as q_ψ, where ψ are the parameters of the distribution q. For q_ψ to approximate $p(\theta|x)$, we learn the distribution parameters ψ using a neural network $F_{(x,\phi)}$ where ϕ are the trainable network weights. Once the network is fully trained and converged, we can state that

$$q_{F(x,\phi)} \sim p(\theta|x). \tag{2}$$

To train $F(x,\phi)$, we minimise the following loss function:

$$L(\phi) = -\sum_{j=1}^{N} \log q_{F(x_j,\phi)}(\theta_j) \tag{3}$$

where: N is the number of samples in the dataset,x_j represents the simulated data points, θ_j are the corresponding parameters drawn from the prior or proposal distribution and $q_{F(x_j,\phi)}(\theta_j)$ is the network's estimated probability of θ_j given x_j.

Once trained, the target posterior can be approximated for all parameters within the prior volume instantaneously. This is referred to as amortization when no additional sampling/training is required for a new set of data, making amortized SBI approaches competitive with state-of-the-art RFS techniques.

[1]This assumes that the data set is contained within the trained parameter distribution

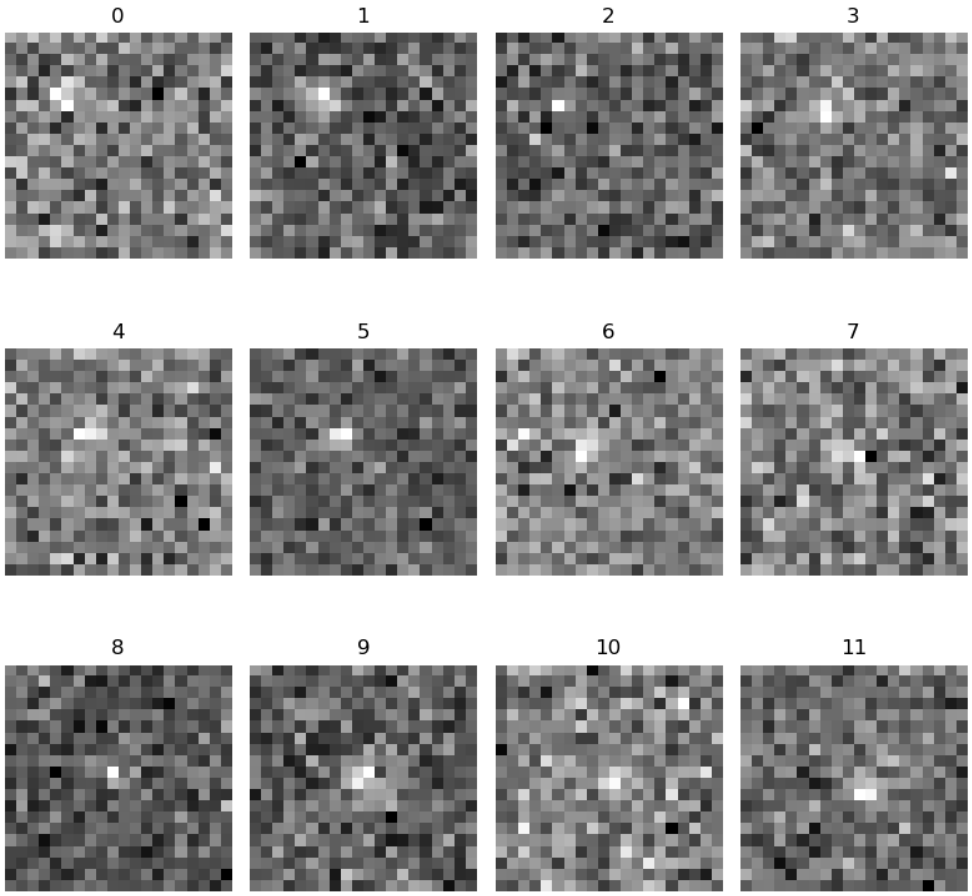

Fig. 1: Figure showing simulations of a time series of observations of an RSO at signal to noise ratios of ~ 3. The RSO is assumed to have a Point Spread Function of ~ 1 pixel. The indices denote the frame sequence with the RSO travelling at a 0.5 pixels rate in both axes per frame observed.

3. DATA SIMULATION

To illustrate the use of amortized SBI, we have created a simple simulation of a low SNR RSO in cis-lunar orbit. We do not intend to be specific as to the type of RSO, orbit or other physical characteristics. This simulation's only purpose is to demonstrate the use of SBI in retrieving the orbital state vector (i.e. position and linear velocity).

We hence simulate an RSO with a 2D Gaussian profile with a 1 pixel Full Width Half Max (FWHM) and a SNR of ~ 3 defined by the peak of the Gaussian profile and the local mean noise amplitude. We assume Normally distributed noise throughout.

The orbital model is assumed to be linear and we define the state space model as a combination of initial position and linear velocity (x, y, vx, vy). We note that non-linear models can easily be implemented but chose to restrict our example to a simple linear case. In figure 1 we show a simulation over 12 consecutive frames with a frame-by-frame RSO velocity of 0.5 pixels in each axis.

We generate two sets of data: A training set and a test set. The training set is used to train $F_{(x,\phi)}$ and consists of parameter and observation pairs (θ, x). We generate these on the fly whilst training the network itself but note that this set can be produced in advance or obtained as by-product of earlier MCMC inverse modelling runs.

The training set is simulated at the same SNR as the test set observation. However, experimenting, we find that it is not necessarily required to train the neural network at precisely the same SNR and higher SNRs reduce the number of epochs the network requires to converge. We attribute this to the Gaussian noise characteristics and following the central limit theorem postulate that a lower SNR in the training data can be compensated with an increase in training data. Finally, the test set is the tested observation. In this case it is another sample simulated from the prior distribution. In future work, we aim to inject RSOs in real data to better increase the realism of the test set.

4. INFERENCE RESULTS

Once our neural network is trained and converged, we can proceed to the inference step. This is a simple step where we provide a new data set, x_0, to the trained SBI inference and obtain its posterior approximation by sampling from the posterior density estimator $q_{F(x,\phi)}$ N times. In effect, we run the trained neural network N iterations with the input x_0 to obtain samples describing our posterior distribution. In figure 2 we show a typical example of the obtained posterior distribution with $N = 1000$ samples. Here, the red lines denote the ground-truth which were well recovered by the approximated posteriors. We should note that very low SNR scenarios require significant training to converge the surrogate distribution q to the correct posterior. However, this training in most cases only needs to occur once with subsequent inference steps being near instantaneous. This offloads the computationally expensive sampling of MCMC techniques to the training phase and allows subsequent real-time inference of the full posterior distribution of the orbital state space model.

5. DISCUSSION

While in this proof of concept paper we only demonstrate the barebone workings of SBI in weak signal RSO detections, we will extend this work to include more realistic training data derived from real observations. This will help quantify the exact convergence statistics in low SNR detections. Likelihood free approaches such as SBI have multiple advantages over more classical approaches. By their very nature, likelihoods do not need to be explicitly specified but can either be marginalised over or specifically derived from simulations. This has interesting applications in areas of SSA that have difficult likelihoods through the interaction of non-Gaussian noise with strongly non-linear models, or in areas where the heterogeneous nature of the likelihood makes a holistic definition difficult, such as data fusion applications.

6. CONCLUSION

In this paper we introduced and demonstrated the use of simulation based inference for the retrieval of orbital state space models in low signal to noise observing conditions. We train our SBI neural network using time series simulations of optical RSO observations. These simulations span a wide range of parameters resulting in an amortized neural posterior estimator. Once converged, our method can approximate the true posterior distribution in real time without the need for further time expensive sampling. In the future we will extend this work beyond this simple proof of concept to include real-world training data.

7. REFERENCES

[1] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo*. CRC Press, May 2011. Google-Books-ID: qfRsAIKZ4rIC.

[2] Daniel Clark, Emmanuel Delande, and Carolin Frueh. Multi-Object Filtering for Space Situational Awareness. 2014.

[3] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, December 2020. arXiv:1911.01429 [cs, stat].

[4] Audrey Cuillery and Francois Le Gland. Track-before-detect on radar image observation with an adaptive auxiliary particle filter. *2019 22th International Conference on Information Fusion (FUSION)*, pages 1–8, 2019.

[5] S.J. Davey, M.G. Rutten, B. Cheung, and B. Cheung. A Comparison of Detection Performance for Several Track-before-Detect Algorithms. *EURASIP Journal on Advances in Signal Processing*, 2008(1):428036, 2007.

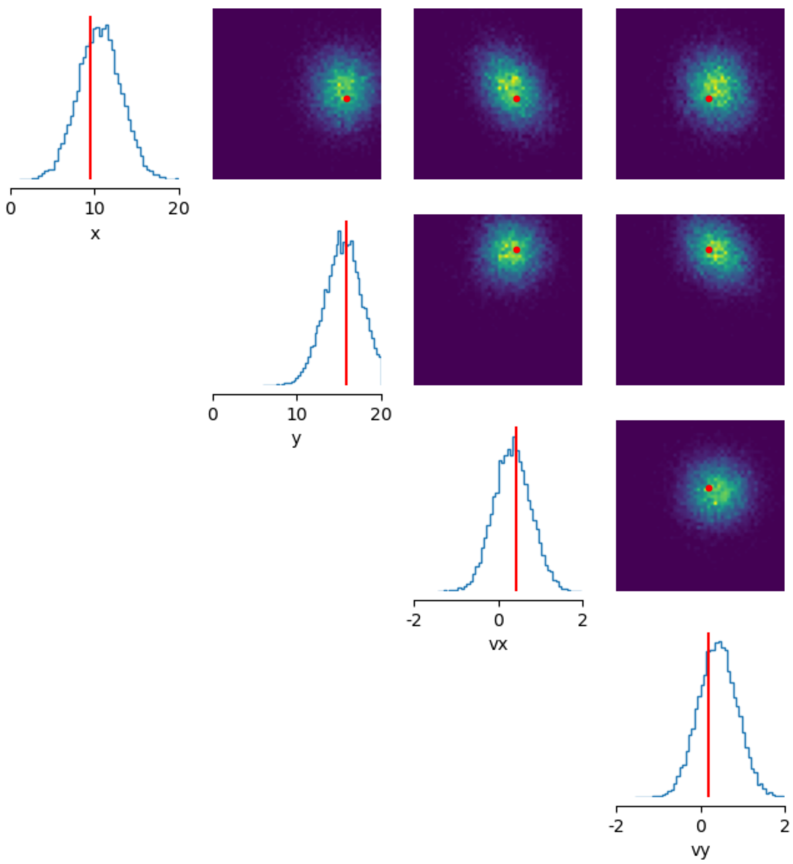

Fig. 2: Showing the approximated posterior distribution q_ψ for a test set example. The red line indicates the true value and the histograms show the marginalised posterior distribution with the coloured heat maps being the conditional posterior distributions. As this is an amortized inference, this result was obtained in ~ 1 second on MacBook Air.

[6] F. Feroz, M. P. Hobson, and M. Bridges. MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. , 398(4):1601–1614, October 2009.
[7] Angel F. Garca-Fernndez, Jason L. Williams, Karl Granstrm, and Lennart Svensson. Poisson Multi-Bernoulli Mixture Filter: Direct Derivation and Implementation. *IEEE Transactions on Aerospace and Electronic Systems*, 54(4):1883–1901, 2016. _eprint: 1703.04264.
[8] Ángel F. García-Fernández, Lennart Svensson, Jason L. Williams, Yuxuan Xia, and Karl Granström. Trajectory multi-Bernoulli filters for multi-target tracking based on sets of trajectories. *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, 00:1–8, 2020.
[9] David S. Greenberg, Marcel Nonnenmacher, and Jakob H. Macke. Automatic Posterior Transformation for Likelihood-Free Inference, May 2019. arXiv:1905.07488 [cs, stat].
[10] Islam Hussein, Matthew P Wilkins, Christopher W T Roscoe, Weston Faber, Suman Chakravorty, and Paul W Schumacher. A High Performance Computing Study of a Scalable FISST-Based Approach to Multi-Target, Multi-Sensor Tracking. 2016.
[11] Liangjie Jia, Peng Rao, Yuke Zhang, Yueqi Su, and Xin Chen. Low-SNR Infrared Point Target Detection and Tracking via Saliency-Guided Double-Stage Particle Filter. *Sensors*, 22(7):2791, 2022.
[12] Mingchao Liang, Thomas Kropfreiter, and Florian Meyer. A BP Method for Track-Before-Detect. *arXiv*, 2023. _eprint: 2307.00765.
[13] Ronald Mahler. Measurement-to-Track Association and Finite-Set Statistics. *arXiv*, 2017. _eprint: 1701.07078.
[14] Stephan Reuter, Benjamin Wilking, Jurgen Wiest, Michael Munz, and Klaus Dietmayer. Real-Time Multi-Object Tracking using Random Finite Sets. *IEEE Transactions on Aerospace and Electronic Systems*, 49(4):2666–2678, 2013.
[15] D.J. Salmond and H. Birch. A particle filter for track-before-detect. *Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148)*, 5:3755–3760 vol.5, 2001.
[16] John Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833–859, December 2006. Publisher: International Society for Bayesian Analysis.
[17] Shahzad Virani and Marcus Holzinger. Real-time Multi-target Detection & Tracking of Space Objects using FiSSt Methods. In *AMOS*, 2020.
[18] Baishen Wei and Brett D. Nener. Multi-Sensor Space Debris Tracking for Space Situational Awareness With Labeled Random Finite Sets. *IEEE Access*, 7:36991–37003, 2019.
[19] Yuxuan Xia, Karl Granström, Lennart Svensson, Ángel F García-Fernández, and Jason L Williams. Multi-Scan Implementation of the Trajectory Poisson Multi-Bernoulli Mixture Filter. *arXiv*, 2019. _eprint: 1912.01748.
[20] Andrew Zammit-Mangion, Matthew Sainsbury-Dale, and Raphaël Huser. Neural methods for amortised inference, 2024.