Model Exploration through Marginal Likelihood Entropy Maximisation

Daniel Jarne Ornia*
University of Oxford
daniel.jarneornia@cs.ox.ac.uk

Joel Dyer*
University of Oxford
joel.dyer@cs.ox.ac.uk

Nicholas Bishop*
University of Oxford
nicholas.bishop@cs.ox.ac.uk

Anisoara Calinescu
University of Oxford

Michael Wooldridge
University of Oxford

Abstract

Simulation modelling offers a flexible approach to constructing high-fidelity synthetic representations of real-world complex systems. The appeal of such models often lies in their ability to facilitate scenario exploration: exploring the different possible futures that could manifest in a complex system. Using simulators for this purpose requires efficient procedures for exploring the range of possible behaviours the simulator can produce. In this paper, we propose and investigate a method to efficiently explore, in an end-to-end parametric manner, the different behaviours that can arise from stochastic, differentiable black-box simulators. Our approach entails maximising the entropy of the marginal likelihood function induced by a trainable proposal distribution over the model's parameter space, computed using direct entropy estimators of the simulated outputs. The method does not require the simulators to have tractable likelihood functions, does not entail building entropy surrogates or instantiating multiple different models, and can be easily parallelised. We provide a proof-of-concept demonstration of the effectiveness of our proposed method on an epidemic simulator commonly used in the literature.

1 Introduction

The growing availability of cheap computational power over recent decades has enabled widespread use of *simulation* modelling to study complex systems across scientific domains such as applied physics [Tompson et al., 2017, Moss et al., 2023], economics [Dyer et al., 2024, Wiese et al., 2024], and epidemiology [Kerr et al., 2021]. Due to their high fidelity, simulators can be used to gain insights into system dynamics, generate synthetic datasets, or identify effective intervention strategies that can be deployed in the real world. However, model construction typically relies on strong structural assumptions, geometric priors and domain expert knowledge. While this is useful, and perhaps necessary, to reduce model complexity, it can lead to reduced understanding and intuition of (parameterised) model behaviour, whilst potentially introducing undesirable modes and edge cases. For this reason, it is useful to develop computational techniques for exploring the behaviour of simulation models, to help practitioners (and users of these models) understand the range of behaviours a simulation model can produce as its parameters are changed.

As a first attempt at model exploration, one may consider the naive approach of forward-simulating using parameters drawn from a maximum entropy distribution. That is, to use parameters drawn from a uniform distribution over the parameter domain. However, it is *not* guaranteed that generating

*Equal contribution

D3S3: Data-driven and Differentiable Simulations, Surrogates, and Solvers @ NeurIPS 2024.

simulations using parameters drawn from a uniform distribution will efficiently reveal the full range of behaviours that a simulator can produce. For example, a diverse set of model behaviours may only be observable through the use of parameters belonging to a small and concentrated region of the parameter space. Correspondingly, large regions of the parameter space may produce relatively similar or indistinguishable outputs. In such cases, a uniform distribution over model parameters samples inefficiently from the full spectrum of behaviours that a model could produce.

In this paper, we propose a method for learning a distribution over simulators parameters that induces the greatest diversity in behaviours. We choose as our measure of diversity the (differential) entropy of the marginal likelihood function induced by this distribution over parameters. We learn this proposal distribution in a variational manner by optimising a Kozachenko-Leonenko entropy estimator [Kozachenko and Leonenko, 1987, Kraskov et al., 2004], constructed from sampled data. Our approach requires that the simulator is differentiable, in the sense that the derivative of simulated samples with respect to the input parameters can be defined and computed via automatic differentiation. We demonstrate the effectiveness of the method for an example simulation model, and discuss the benefits and caveats of the proposed method.

2 Simulators and Scenario Exploration

In this work, we consider a stochastic simulator characterised by a tuple $(f, p_{\mathbf{u}})$. The function $f : \Theta \times \mathcal{U} \to \mathcal{X}$ describes how parameters $\theta \in \Theta$ and noise $\mathbf{u} \in \mathcal{U}$ are mapped to simulation outputs $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^m$, whilst $p_{\mathbf{u}}$ describes the noise distribution over $\mathcal{U}$ implicitly defined by the simulator. Forward-simulation with parameter θ corresponds to first sampling noise $\mathbf{u} \sim p_{\mathbf{u}}$ and then evaluating $f(\theta, \mathbf{u})$. Note that a simulator implicitly defines a likelihood function $p(\mathbf{x} \mid \theta)$. For a proposal distribution q over the parameter space Θ, the marginal likelihood function of the simulator is

$$p(\mathbf{x} \mid q) = \int_{\Theta} p(\mathbf{x} \mid \theta) q(\theta) d\theta.$$

Due to the complex nature of stochastic simulators, we assume that $p(\mathbf{x} \mid \theta)$, and thus $p(\mathbf{x} \mid q)$, cannot be tractably evaluated. In contrast, we assume throughout that the simulator is differentiable: we assume that $\frac{\partial f(\theta, \mathbf{u})}{\partial \theta}$ is well-defined at all points $(\theta, \mathbf{u})$, or that suitable "surrogate gradients" [see e.g. Maheswaranathan et al., 2019] can be constructed when this requirement is not satisfied.

Given a stochastic simulator $(f, p_{\mathbf{u}})$, our goal is to identify a proposal distribution q produces a diverse range of outputs from the simulator. In other words, we seek a proposal distribution q that ensures all observable outputs can be sampled with relatively high probability. This corresponds to ensuring that the marginal likelihood distribution $p(\mathbf{x} \mid q)$ produces a diverse range of samples.

Note that entropy may be viewed as a natural measure of information diversity. Thus, maximising the entropy of the marginal likelihood distribution associated with q forms a natural objective:

$$\arg\min_{q} \int_{\mathcal{X}} p(\mathbf{x} \mid q) \log p(\mathbf{x} \mid q) d\mathbf{x}. \tag{1}$$

As discussed, the marginal likelihood function $p(\mathbf{x} \mid q)$ cannot be tractably evaluated in general, preventing direct empirical estimation of the entropy objective in Problem (1) through standard Monte Carlo methods. Moreover, when the dimensionality of the output space $\mathcal{X}$ is high, optimising over all possible proposal distributions rapidly becomes intractable. To address these issues, we propose a variational optimisation approach based on the Kozachenko-Leonenko entropy estimator [Kozachenko and Leonenko, 1987, Kraskov et al., 2004] in the next section.

3 Differentiable Entropy Estimation

In order to make Problem (1) tractable, we first restrict q to a variational family $\mathcal{Q} = \{q_\phi \mid \phi \in \Phi\}$ parameterised by $\phi \in \Phi$. We further assume that q_ϕ is differentiable with respect to ϕ. Many rich variational families, such as those associated with normalising flows, satisfy this criterion. Note that identifying the optimal proposal distribution $q^\star$ now corresponds to identifying the optimal parameter values $\phi^\star$. Given the differentiability of the simulator $(f, p_{\mathbf{u}})$ it is natural to consider a stochastic gradient descent scheme to learn $\phi^\star$:

$$\phi_{t+1} \leftarrow \texttt{Update}\left(\phi_t, g(\mathcal{D}_{\phi_t}), \eta\right) \tag{2}$$

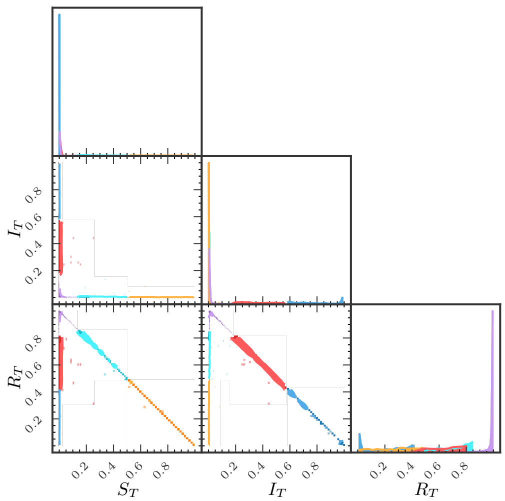

Figure 1: The marginal likelihood induced by a uniform distribution over model parameters for the SIR agent-based model. Colours show k-means clustering results.

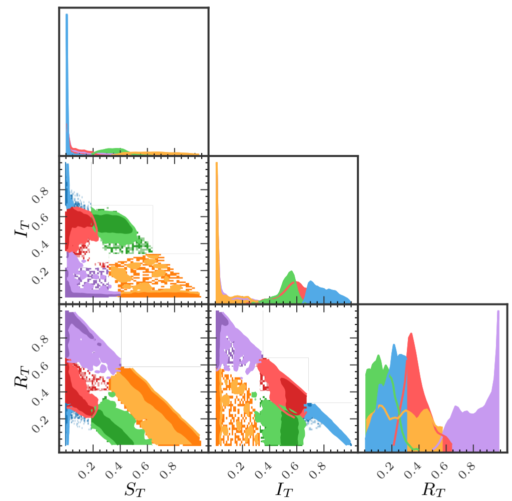

Figure 2: The maximum entropy marginal likelihood uncovered with our optimisation procedure for the SIR agent-based model. Colours show k-means clustering results.

where g is a Monte-Carlo gradient estimate computed using the stochastic batch $\mathcal{D}_{\phi_t} = (\mathbf{x}_i)_{i=1}^n$ sampled from the marginal likelihood $p(\cdot \mid q_{\phi_t})$ through forward-simulation, η is the learning rate, and `Update` is some procedure for updating the parameters of the proposal distribution using gradient information. However, since the marginal likelihood $p(\mathbf{x} \mid q_{\phi_t})$ cannot be efficiently evaluated, constructing a Monte-Carlo gradient estimate g is non-trivial.

To address this, we propose the use of the Kozachenko-Leonenko (KL) estimator [Kozachenko and Leonenko, 1987, Kraskov et al., 2004]. Let $\mathbf{x}_i^k$ denote the kth nearest neighbour by Euclidean distance to output $\mathbf{x}_i$ in the batch $\mathcal{D}_{\phi_t}$. The Kozachenko-Leonenko entropy estimator $\hat{H}(\mathcal{D}_{\phi_t})$ is given by

$$\hat{H}(\mathcal{D}_{\phi_t}) := -\psi(k) + \psi(n) + \log(v_m) + \frac{m}{n}\sum_{i=1}^{n} \log\left(2\|\mathbf{x}_i - \mathbf{x}_i^k\|\right), \tag{3}$$

where ψ denotes the digamma function and v_m denotes the volume of the Euclidean unit ball in $\mathbb{R}^m$. Note that $\hat{H}(\mathcal{D}_{\phi_t})$ is differentiable with respect to ϕ when both the variational family $\mathcal{Q}$ and the simulator $(f, p_{\mathbf{u}})$ are differentiable with respect to their parameters. By substituting $\hat{H}(\mathcal{D}_{\phi_t})$ into Problem (1) and applying a standard optimisation algorithm, e.g., AdamW [Loshchilov and Hutter, 2019], as the `Update` operator, we obtain a practical training scheme for ϕ using $\nabla_\phi \hat{H}(\mathcal{D}_\phi)|_{\phi_t}$.

4 Experiments

Next, we present a proof-of-concept demonstration of our methodology using an agent-based susceptible-infected-recovered (SIR) epidemic simulator. We define the variational family $\mathcal{Q}$ using a normalising flow which is trained via the gradient scheme described in Section 3. Full details of the flow architecture and simulator used in our experiments is provided in the supplementary material.

SIR Epidemic Simulation. We consider an SIR epidemic simulator consisting of a population of individuals connected through a graph structure, amongst whom a simulated "virus" spreads. At each time step $t = 1, \ldots, T$, each susceptible agent i becomes infected with probability $1 - (1 - \beta)^{n_i}$, where n_i denotes the number of infected agents that neighbour i. Agents that are currently infected recover spontaneously at each time step with probability γ. We use S_t, I_t and R_t to denote the relative proportion of susceptible, infected and recovered agents on time step t. The model has three parameters, $\theta = (I_0, \beta, \gamma) \in [0, 1]^3$. A simulation output $\mathbf{x} = (S_T, I_T, R_T)$ describes the final proportion of infected, susceptible and recovered individuals. Note that this model involves discrete randomness to determine the transitions of agents between discrete states. However, the model can be made differentiable by adopting, for example, straight-through or Gumbel-Softmax surrogate gradients [see, e.g., Bengio et al., 2013, Jang et al., 2016].

Figures 1 and 2 show 10^4 samples from the marginal likelihood function $p(\mathbf{x} \mid q)$ when q is, respectively, a uniform distribution and the learned, entropy-maximising distribution over the parameter space, obtained using the procedure we describe in Sections 2 and 3. Colours correspond to the clustering labels found from running a k-means clustering [Lloyd, 1982] with 5 clusters. As is evident from these plots, our proposed method for exploring the full range of behaviours the model is able to generate has successfully teased out model behaviours that are indistinct or altogether unseen when q is taken to be a uniform distribution. The most visually striking example is the cluster coloured green in Figure 2, which captures model dynamics that lead to a final state consisting of a moderate proportion of both susceptible and infected individuals, and low numbers of recovered individuals; it is evident that such behaviour is entirely missed in Figure 1. Density plots for the obtained maximum entropy prior distribution are included in Appendix A.3.

5 Related Work

Entropy Estimation. Entropy estimation is a longstanding statistical problem, and a vast range of estimators have been proposed. These include plug-in estimators [Györfi and van der Meulen, 1987], methods based on sample spacings [Tarasenko, 1968], and nearest neighbour approaches [Berrett et al., 2019, Tsybakov and van der Meulen, 1996], such as the Kozachenko-Leonenko estimator [Kozachenko and Leonenko, 1987]. We chose to adopt the Kozachenko-Leonenko estimator since it is differentiable and, in contrast to many plug-in estimators, does not explicitly require density estimation which may be costly as the dimensionality of the output space grows.

Exploration in RL. Our approach is closely related to maximum entropy exploration for reinforcement learning (RL) [Hazan et al., 2019, Tiapkin et al., 2023, Zahavy et al., 2024], wherein the goal is to learn policies that induce highly uniform distributions over trajectories or state-action pairs. More broadly, our work shares its motivation with reward-free exploration methods, such as count-based exploration [Tang et al., 2017, Bellemare et al., 2016] and intrinsic RL [Chentanez et al., 2004, Achiam and Sastry, 2017]. Whereas the aforementioned methods aim to learn randomised policies for finite Markov decision processes, we aim to learned randomised parameters for arbitrary differentiable stochastic simulators with potentially infinite output spaces.

Minimum Energy Design. Our problem setting is also closely related to both minimum energy [Joseph et al., 2015] and maximin [Li et al., 2020] design, which attempt to identify a finite set of design points that characterise the response surface of a deterministic simulator. In contrast, we consider *stochastic* simulators, and aim to produce a *distribution* over simulation parameters that induces a uniform distribution over possible outputs. Whilst the aforementioned design problems require computationally expensive approaches such as genetic algorithms and simulated annealing, we obtain a variational problem that can be efficiently solved via gradient ascent.

6 Conclusions and Future Directions

We have presented a new model exploration framework for differentiable simulators based on maximisation of the differential entropy. As demonstrated empirically, our methodology can help reveal simulator behaviours frequently missed by uniform sampling of the parameter space. With that being said, there are many promising directions for future work. As entropy estimation generally scales poorly with dimension, we chose to focus on low dimensional summary statistics in our experiments. Generalising our approach to high dimensional outputs such as time series presents an interesting direction for future work. Lastly note that a modeller may be interested in exploring the distributional or aggregate quantities such as the average output associated with a given parameter value. Generalising our approach to such a setting forms another interesting direction for future research.

Acknowledgments and Disclosure of Funding

JD, NB, DJ, AC, and MW acknowledge funding from a UKRI AI World Leading Researcher Fellowship awarded to Wooldridge (grant EP/W002949/1). MW and AC also acknowledge funding from Trustworthy AI - Integrating Learning, Optimisation and Reasoning (TAILOR), a project funded by European Union Horizon2020 research and innovation program under Grant Agreement 952215. DJ, AC and MW acknowledge in part a grant from the Alan Turing Institute, London.

References

Joshua Achiam and Shankar Sastry. Surprise-based intrinsic motivation for deep reinforcement learning. *arXiv preprint arXiv:1703.01732*, 2017.

Marc Bellemare, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, and Remi Munos. Unifying count-based exploration and intrinsic motivation. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016.

Yoshua Bengio, Nicholas Léonard, and Aaron Courville. Estimating or propagating gradients through stochastic neurons for conditional computation. *arXiv preprint arXiv:1308.3432*, 2013.

Thomas B Berrett, Richard J Samworth, and Ming Yuan. Efficient multivariate entropy estimation via k-nearest neighbour distances. *The Annals of Statistics*, 47(1):288–318, 2019.

Nuttapong Chentanez, Andrew Barto, and Satinder Singh. Intrinsically motivated reinforcement learning. In L. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 17. MIT Press, 2004.

Joel Dyer, Patrick Cannon, J Doyne Farmer, and Sebastian M Schmon. Black-box bayesian inference for agent-based models. *Journal of Economic Dynamics and Control*, 161:104827, 2024.

László Györfi and Edward C. van der Meulen. Density-free convergence properties of various estimators of entropy. *Computational Statistics & Data Analysis*, 5(4):425–436, 1987. ISSN 0167-9473.

Elad Hazan, Sham Kakade, Karan Singh, and Abby Van Soest. Provably efficient maximum entropy exploration. In *International Conference on Machine Learning*, pages 2681–2691. PMLR, 2019.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. In *International Conference on Learning Representations*, 2016.

V. Roshan Joseph, Tirthankar Dasgupta, Rui Tuo, and C. F. Jeff Wu. Sequential exploration of complex surfaces using minimum energy designs. *Technometrics*, 57(1):64–74, 2015. ISSN 00401706, 15372723.

Cliff C Kerr, Robyn M Stuart, Dina Mistry, Romesh G Abeysuriya, Katherine Rosenfeld, Gregory R Hart, Rafael C Núñez, Jamie A Cohen, Prashanth Selvaraj, Brittany Hagedorn, et al. Covasim: an agent-based model of covid-19 dynamics and interventions. *PLOS Computational Biology*, 17(7): e1009149, 2021.

Lyudmyla F Kozachenko and Nikolai N Leonenko. Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii*, 23(2):9–16, 1987.

Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 69(6):066138, 2004.

Wenlong Li, Min-Qian Liu, and Boxin Tang. A method of constructing maximin distance designs. *Biometrika*, 108(4):845–855, 10 2020. ISSN 0006-3444. doi: 10.1093/biomet/asaa089.

Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2): 129–137, 1982.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*. PMLR, 2019.

Niru Maheswaranathan, Luke Metz, George Tucker, Dami Choi, and Jascha Sohl-Dickstein. Guided evolutionary strategies: Augmenting random search with surrogate gradients. In *International Conference on Machine Learning*, pages 4264–4273. PMLR, 2019.

Guy Moss, Vjeran Višnjević, Olaf Eisen, Falk M Oraschewski, Cornelius Schröder, Jakob H Macke, and Reinhard Drews. Simulation-based inference of surface accumulation and basal melt rates of an antarctic ice shelf from isochronal layers. *arXiv preprint arXiv:2312.02997*, 2023.

George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. *Advances in neural information processing systems*, 30, 2017.

Arnau Quera-Bofarull, Joel Dyer, Anisoara Calinescu, J Doyne Farmer, and Michael Wooldridge. Blackbirds: Black-box inference for differentiable simulators. *Journal of Open Source Software*, 8 (89), 2023.

Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pages 1530–1538. PMLR, 2015.

Vincent Stimper, David Liu, Andrew Campbell, Vincent Berenz, Lukas Ryll, Bernhard Schölkopf, and José Miguel Hernández-Lobato. normflows: A pytorch package for normalizing flows. *Journal of Open Source Software*, 8(86):5361, 2023.

Haoran Tang, Rein Houthooft, Davis Foote, Adam Stooke, OpenAI Xi Chen, Yan Duan, John Schulman, Filip DeTurck, and Pieter Abbeel. # exploration: A study of count-based exploration for deep reinforcement learning. *Advances in neural information processing systems*, 30, 2017.

F.P. Tarasenko. On the evaluation of an unknown probability density function, the direct estimation of the entropy from independent observations of a continuous random variable, and the distribution-free entropy test of goodness-of-fit. *Proceedings of the IEEE*, 56(11):2052–2053, 1968. doi: 10.1109/PROC.1968.6784.

Daniil Tiapkin, Denis Belomestny, Daniele Calandriello, Eric Moulines, Remi Munos, Alexey Naumov, Pierre Perrault, Yunhao Tang, Michal Valko, and Pierre Menard. Fast rates for maximum entropy exploration. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 34161–34221. PMLR, 23–29 Jul 2023.

Jonathan Tompson, Kristofer Schlachter, Pablo Sprechmann, and Ken Perlin. Accelerating eulerian fluid simulation with convolutional networks. In *International conference on machine learning*, pages 3424–3433. PMLR, 2017.

A. B. Tsybakov and E. C. van der Meulen. Root-n consistent estimators of entropy for densities with unbounded support. *Scandinavian Journal of Statistics*, 23(1):75–83, 1996. ISSN 03036898, 14679469.

Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393 (6684):440–442, 1998.

Samuel Wiese, Jagoda Kaszowska-Mojsa, Joel Dyer, Jose Moran, Marco Pangallo, Francois Lafond, John Muellbauer, Anisoara Calinescu, and J Doyne Farmer. Forecasting macroeconomic dynamics using a calibrated data-driven agent-based model. *arXiv preprint arXiv:2409.18760*, 2024.

Tom Zahavy, Brendan O'Donoghue, Guillaume Desjardins, and Satinder Singh. Reward is enough for convex mdps. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, Red Hook, NY, USA, 2024. Curran Associates Inc. ISBN 9781713845393.

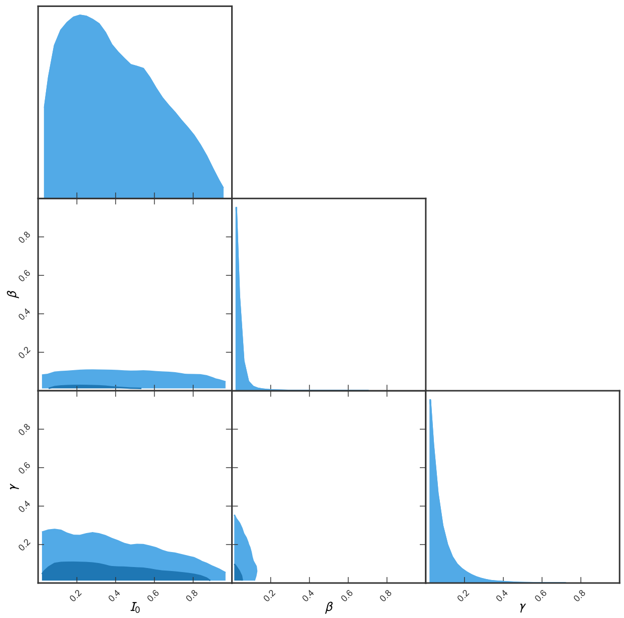

Figure 3: Prior density of parameters learned through entropy maximization.

A Further experimental details

A.1 Further details on the SIR agent-based model

In our experiments, we make use of the differentiable implementation of an agent-based SIR model appearing in the `BlackBIRDS` software package [Quera-Bofarull et al., 2023]. In this model, all agents have probability I_0 of beginning in a state of infection, else they begin the simulation as susceptible agents. The gradients of all Bernoulli random variables used to determine the discrete transitions between agent states are constructed using the Gumbel-Softmax gradient trick [Jang et al., 2016]. In every simulation, we simulate 200 agents for $T = 1000$ time steps on a Watts-Strogatz random network [Watts and Strogatz, 1998] with 10 links per node and rewiring probability 0.1.

A.2 Variational proposal distribution details

As the variational proposal distribution in the entropy maximisation scheme, we use a normalising flow [Rezende and Mohamed, 2015] consisting of 5 transforms, each of which consists of a Masked Autoregressive layer [Papamakarios et al., 2017] with 2 blocks with 20 hidden features, followed by a permutation layer. The flow was built using the `normflows` package [Stimper et al., 2023], and trained using the AdamW optimiser [Loshchilov and Hutter, 2019] with a learning rate of 10^{-3} for a maximum of 10^3 epochs. We cease training if the loss does not decrease for 100 consecutive epochs.

A.3 Learned Priors

We include here the prior density plots for the trained proposal distributions when maximising model output entropy. Figure 3 shows the parameter densities obtained by maximising entropy of the model output.