

---

# Coping With Simulators That Don't Always Return

---

**Andrew Warrington**  
University of Oxford  
andreww@robots.ox.ac.uk

**Saeid Naderiparizi**  
University of British Columbia  
saeidnp@cs.ubc.ca

**Frank Wood**  
University of British Columbia  
fwood@cs.ubc.ca

## Abstract

Deterministic models are approximations of reality that are easy to interpret and often easier to build than stochastic alternatives. Unfortunately, as nature is capricious, observational data can never be fully explained by deterministic models in practice. Observation and process noise need to be added to adapt deterministic models to behave stochastically, such that they are capable of explaining and extrapolating from noisy data. We investigate and address computational inefficiencies that arise from adding process noise to deterministic simulators that fail to return for certain inputs; a property we describe as “brittle.” We show how to train a conditional normalizing flow to propose perturbations such that the simulator succeeds with high probability, increasing computational efficiency.

## 1 Introduction

In order to compensate for epistemic uncertainty due to modelling approximations and unmodelled aleatoric uncertainty, deterministic simulators are often “converted” to “stochastic” simulators by perturbing the state at each time step. In practice this allows the simulator to explain the variability observed in real data without requiring excessive observation noise. Such models are more resilient to misspecification, are capable of providing uncertainty estimates, and provide better inferences in general [Møller et al., 2011; Saari- nen et al., 2008; Lv et al., 2008; Pimblott and LaVerne, 1990; Renard et al., 2013].

Often, state-independent Gaussian noise with heuris-

tically tuned variance is used to perturb the state [Ad- hikari and Agrawal, 2013; Brockwell and Davis, 2016; Fox, 1997; Reddy and Clinton, 2016; Du and Sam, 2006; Allen, 2017; Mbalawata et al., 2013]. However, naively adding noise to the state will, in many ap- plications, render the perturbed input state “invalid,” where invalid states cause the simulator to raise an exception and not return a value [Razavi et al., 2019; Lucas et al., 2013; Sheikholeslami et al., 2019]. We formally define failure by extending the possible output of the simulator to include  $\perp$  (read as “bottom”) de- noting simulator failure. The principal contribution of this paper is a technique for avoiding invalid states by choosing perturbations that minimize the failure rate. The technique we develop results in a reduction in sim- ulator failures, while maintaining the original model.

Examples of failure modes include ordinary differen- tial equation (ODE) solvers not converging to the re- quired tolerance in the allocated time, or, the per- turbed state entering into an unhandled configura- tion, such as solid bodies intersecting. Establishing the state-perturbation pairs that cause failure is non- trivial. Hence, the simulation artifact can be sensitive to seemingly inconsequential alterations to the state – a property we describe as “brittle.” Failures waste computational resources and reduce the diversity of simulations for a finite sample budget, for instance, when used as the proposal in sequential Monte Carlo. As such, we wish to learn a proposal over perturbations such that the simulator exits with high probability, but renders the joint distribution unchanged.

We proceed by framing sampling from brittle simula- tors as rejection samplers, then seek to eliminate rejec- tions by estimating the state-dependent density over perturbations that do not induce failure. We then demonstrate that using this learned proposal yields lower variance results when used in posterior infer- ence with a fixed sample budget, such as pseudo- marginal evidence estimates produced by sequential Monte Carlo sweeps. Source code for reproduction of figures and results in this paper is available at <https://github.com/plai-group/stdr>.

## 2 Background

### 2.1 Smoothing Deterministic Models

Deterministic simulators are often stochastically perturbed to increase the diversity of the achievable simulations and to fit data more effectively. The most widespread example of this is perturbing linear dynamical systems with Gaussian noise at each timestep. The design of the system is such that the distribution over state at each point in time is Gaussian distributed. However, the simplistic dynamics of such a system may be insufficient for simulating more complex systems. Examples of such systems are: stochastic models of neural dynamics [Fox, 1997; Coutin et al., 2018; Goldwyn and Shea-Brown, 2011; Saarinen et al., 2008], econometrics [Lopes and Tsay, 2011], epidemiology [Allen, 2017] and mobile robotics [Thrun et al., 2001; Fallon et al., 2012]. In these examples, the simulator state is perturbed with noise drawn from a distribution and is iterated using the simulator to create discrete approximations of the distribution over state as a function of time.

### 2.2 Simulator Failure

As simulators become more complex, guaranteeing the simulator will not fail for perturbed inputs becomes more difficult, and individual function evaluations become more expensive. Lucas et al. [2013] and Edwards et al. [2011] establish the sensitivity of earth science models to global parameter values by building a discriminative classifier for parameters that induce failure. Sheikholeslami et al. [2019] take an alternative approach instead treating simulator failure as an imputation problem, fitting a function regressor to predict the outcome of the failed experiment given the neighboring experiments that successfully terminated. However these methods are limited by the lack of clear probabilistic interpretation in terms of the originally specified joint distribution in time series models, their ability to scale to high dimensions, and their applicability to state-space models.

### 2.3 State-space Inference and Model Selection

Probabilistic models are ultimately deployed to make inferences about the world. Hence the goal is to be able to recover distributions over unobserved states, predict future states and learn unknown parameters of the model from data. Posterior state-space inference refers to the task of recovering the distribution  $p_{\mathcal{M}}(\mathbf{x}_{0:T}|\mathbf{y}_{1:T})$ , where  $\mathbf{x}_{0:T}$  are the latent states,  $\mathbf{y}_{1:T}$  are the observed data, and  $\mathcal{M}$  denotes the model if multiple different models are available. Inference in

---

### Algorithm 1 Sequential Monte Carlo

---

```

1: procedure SMC( $p_{\mathcal{M}}(\mathbf{x}_0)$ ,  $\bar{p}_{\mathcal{M}}(\mathbf{x}_t|\mathbf{x}_{t-1})$ ,  $\mathbf{y}_{1:T}$ ,
    $p_{\mathcal{M}}(\mathbf{y}_t|\mathbf{x}_t)$ ,  $N$ )
2:   for  $n = 1 : N$  do
3:      $\mathbf{x}_0^{(n)} \sim p_{\mathcal{M}}(\mathbf{x}_0)$   $\triangleright$  Initialize from prior.
4:      $L_{\mathcal{M}} \leftarrow 0$   $\triangleright$  Track log-evidence
5:     for  $t = 1 : T$  do
6:       for  $n = 1 : N$  do
7:          $\tilde{\mathbf{x}}_t^{(n)} \sim \bar{p}_{\mathcal{M}}(\mathbf{x}_t|\mathbf{x}_{t-1}^{(n)})$   $\triangleright$  Alg 2.
8:          $w_t^{(n)} \leftarrow p_{\mathcal{M}}(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(n)})$   $\triangleright$  Score particle.
9:       for  $n = 1 : N$  do  $\triangleright$  Normalize weights.
10:         $W_t^{(n)} \leftarrow w_t^{(n)} / \sum_{i=1}^N w_t^{(i)}$ 
11:       for  $n = 1 : N$  do  $\triangleright$  Apply resampling.
12:         $a_t^{(n)} \sim \text{Discrete}(\mathbf{W}_t)$ 
13:         $\mathbf{x}_t^{(n)} \leftarrow \tilde{\mathbf{x}}_t^{(a_t^{(n)})}$ 
14:         $L_{\mathcal{M}} \leftarrow L_{\mathcal{M}} + \log\left(\frac{1}{N} \sum_{i=1}^N w_t^{(i)}\right)$ 
15:   return  $\mathbf{x}_{0:T}^{(1:N)}$ ,  $a_{1:T}^{(1:N)}$ ,  $L_{\mathcal{M}}$ 
16: end procedure

```

---

Gaussian perturbed linear dynamical systems can be performed using techniques such as Kalman smoothing [Kalman et al., 1960], however, the restrictions on such techniques limit their applicability to complex simulators, and so numerical methods are often used in practice.

A common method for performing inference in complex, simulation based models is sequential Monte Carlo (SMC) [Doucet et al., 2001]. The basic algorithm for SMC is shown in Algorithm 1, where  $p_{\mathcal{M}}(\mathbf{x}_0)$  is the prior over initial state,  $\bar{p}_{\mathcal{M}}(\mathbf{x}_t|\mathbf{x}_{t-1})$  is the dynamics model, or simulator,  $p_{\mathcal{M}}(\mathbf{y}_t|\mathbf{x}_t)$  is the likelihood, defining the relationship between latent states and observed data, and  $N$  is the number of particles used. On a high level, SMC produces a discrete approximation of the target distribution by iterating particles through the simulator, and then preferentially continuing those simulations that “explain” the observed data well. While a detailed understanding of particle filtering is not required, the core observation required for this work is that the likelihood of failed simulations is defined as zero:  $p(\mathbf{y}_t|\mathbf{x}_t = \perp) := 0$ , and hence are rejected with certainty.

Posterior inference pipelines often also provide estimates of the model evidence,  $p_{\mathcal{M}}(\mathbf{y}_{1:T})$ . SMC provides such an estimate, referred to as a pseudo-marginal evidence, denoted in Algorithm 1 as  $L_{\mathcal{M}}$ . This pseudo-marginal evidence is calculated (in log space) as the sum of the expected value of the unnormalized importance weights (Algorithm 1, Lines 8 and 14). This evidence can be combined with the prior probability

of each model via Bayes rule to estimate the posterior probability of the model (up to a normalizing constant) [MacKay, 2003]. These posteriors can be compared to perform Bayesian model selection, where the model with the highest posterior is selected and used to perform inference. This is often referred to as marginal maximum *a posteriori* parameter estimation (or model selection) [Doucet et al., 2002; Kantas et al., 2015]. Recent work investigates model selection using approximate, likelihood-free inference techniques [Papamakarios et al., 2019; Lueckmann et al., 2019], however, we do not consider these methods here, instead focusing on mitigating computational inefficiencies arising directly from simulator failure.

### 3 Methodology

We consider deterministic models, expressed as simulators, describing the time-evolution of a state  $\mathbf{x}_t \in \mathcal{X}$ , where we denote application of the simulator iterating the state as  $\mathbf{x}_t \leftarrow f(\mathbf{x}_{t-1})$ . A stochastic, additive perturbation to state, denoted  $\mathbf{z}_t \in \mathcal{X}$ , is applied to induce a distribution over states. The distribution from which this perturbation is sampled is denoted  $p(\mathbf{z}_t|\mathbf{x}_{t-1})$ , although, in practice, this distribution is often state independent. The iterated state is then calculated as  $\mathbf{x}_t \leftarrow f(\mathbf{x}_{t-1} + \mathbf{z}_t)$ .

However, we consider the case where the simulator can fail for “invalid” inputs, denoted by a return value of  $\perp$ . Hence the complete definition of  $f$  is  $f : \mathcal{X} \rightarrow \{\mathcal{X}, \perp\}$ . The region of valid inputs is denoted as  $\mathcal{X}_A \subset \mathcal{X}$ , and the region of invalid inputs as  $\mathcal{X}_R \subset \mathcal{X}$ , such that  $\mathcal{X}_A \sqcup \mathcal{X}_R = \mathcal{X}$ , where the boundary between these regions is unknown. Over the whole support,  $f$  defines a many-to-one function, as  $\mathcal{X}_R$  maps to  $\perp$ . However, the algorithm we derive only requires that  $f$  is one-to-one in the accepted region. This is not uncommon in real simulators, and is satisfied by, for example, ODE models. We define the random variable  $A_t \in \{0, 1\}$  to denote whether the state-perturbation pair does not yield simulator failure and is “accepted.”

We define the iteration of perturbed deterministic simulator as a rejection sampler, with a well-defined target distribution (§3.1). We use this definition and the assumptions on  $f$  to show that we can target the same distribution by learning the state-conditional density of perturbations, conditioned on acceptance (§3.2). We train an autoregressive flow to fit this density (§3.3), and describe how this can be used in inference, highlighting the ease with which it can be inserted into a particle filter (§3.4). We empirically show that using this learned proposal distribution in place of the original proposal improves the performance of particle-based state-space inference methods (§4).

---

**Algorithm 2** Iterate brittle simulator,  $\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1})$ .

---

```

1: procedure ITERATESIMULATOR( $f, \mathbf{x}_{t-1}$ )
2:    $\mathbf{x}_t \leftarrow \perp$ 
3:   while  $\mathbf{x}_t == \perp$  do
4:     if  $q_\phi$  is trained then
5:        $\mathbf{z}_t \sim q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$   $\triangleright$  Perturb under  $q_\phi$ .
6:     else
7:        $\mathbf{z}_t \sim p(\mathbf{z}_t|\mathbf{x}_{t-1})$   $\triangleright$  Perturb under  $p$ .
8:      $\mathbf{x}_t \leftarrow f(\mathbf{x}_{t-1} + \mathbf{z}_t)$   $\triangleright$  Iterate simulator.
9:   return  $\mathbf{x}_t$ 
10: end procedure

```

---

#### 3.1 Brittle Simulators as Rejection Samplers

The naive approach to sampling from the perturbed system, shown in Algorithm 2, is to repeatedly sample from the proposal distribution and evaluate  $f$  until the simulator successfully exits. This procedure defines  $A_t = \mathbb{I}[f(\mathbf{x}_{t-1} + \mathbf{z}_t) \neq \perp]$ ,  $\mathbf{z}_t \sim p(\mathbf{z}_t|\mathbf{x}_{t-1})$ , i.e. successfully iterated samples are accepted with certainty. This incurs significant wasted computation as the simulator must be called repeatedly, with failed iterations being discarded. The objective of this work is to derive a more efficient sampling mechanism.

We begin by establishing Algorithm 2 as a rejection sampler, targeting the distribution over successfully iterated states. This reasoning is illustrated in Figure 1. The behavior of  $f$  and the distribution  $p(\mathbf{z}_t|\mathbf{x}_{t-1})$  implicitly define a distribution over successfully iterated states. We denote this “target” distribution as  $\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1}) = p(\mathbf{x}_t|\mathbf{x}_{t-1}, A_t = 1)$ , where the bar indicates that the sample was accepted, and hence places no probability mass on failures. Note there is no bar on  $p(\mathbf{z}_t|\mathbf{x}_{t-1})$ , indicating that it is defined before the accept/reject behaviors of  $f$  and hence probability mass may be placed on regions that yield failure. The functional form of  $\bar{p}$  is unavailable, and the density cannot be evaluated for any input value.

The existence of  $\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1})$  implies the existence of a second distribution: the distribution over *accepted* perturbations, denoted  $\bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1})$ . Note that this distribution is also conditioned on acceptance under the chosen simulator, indicated by the presence of a bar. We assume  $f$  is one-to-one in the accepted region, and so the change of variables rule can be applied to directly relate this to  $\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1})$ . Under our initial algorithm for sampling from a brittle simulator we can therefore write the following identity:

$$\bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1}) = \begin{cases} \frac{1}{M_p} p(\mathbf{z}_t|\mathbf{x}_{t-1}), & \text{if } f(\mathbf{x}_{t-1} + \mathbf{z}_t) \neq \perp \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where the normalizing constant  $M_p$  is the acceptance

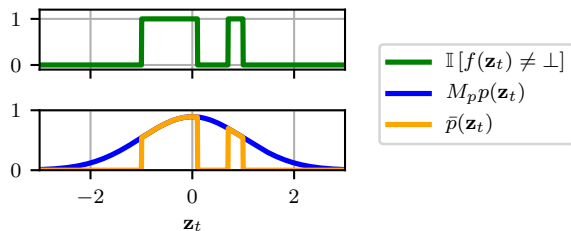


Figure 1: Graphical representation of how a brittle deterministic simulator acts as a rejection sampler, targeting  $\bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1})$ . We set  $\mathbf{x}_t = 0$  for clarity. The simulator,  $f(\mathbf{z}_t)$ , returns  $\perp$  for unknown input regions, shown in green. The proposal over  $\mathbf{z}_t$  is shown in blue. The target distribution,  $\bar{p}(\mathbf{z}_t)$ , shown in orange, is implicitly defined as  $\bar{p}(\mathbf{z}_t) = \frac{1}{M_p} p(\mathbf{z}_t) \mathbb{I}[f(\mathbf{z}_t) \neq \perp]$ , where  $M_p$  is the normalizing constant from  $\bar{p}$ , equal to the acceptance rate. Accordingly, the proposal distribution, scaled by  $M_p$ , is *exactly* equal to  $\bar{p}(\mathbf{z}_t)$  in the accepted region. Algorithm 2 therefore implicitly constructs a rejection sampler, where the acceptance criterion reduces to  $\mathbb{I}[f(\mathbf{z}_t) \neq \perp]$ , without needing to specify any additional scaling constants.

rate under  $p$ . (1) indicates accepting with certainty perturbations that exit successfully can be seen as proportionally shifting mass from regions of  $p$  where the simulator fails to regions where it does not. We exploit this definition to learn an efficient proposal.

### 3.2 Change of Variable in Brittle Simulator

We now derive how we can learn the proposal distribution, denoted  $q_\phi$  and parameterized by  $\phi$ , to replace  $p$ , such that the acceptance rate under  $q_\phi$  (denoted  $M_{q_\phi}$ ) tends towards unity, minimizing wasted computation. We denote  $q_\phi$  as the proposal we train, which, coupled with the simulator, implicitly defines a proposal over accepted samples, denoted  $\bar{q}_\phi$ .

Expressing this mathematically, we wish to minimize the distance between joint distribution implicitly specified over *accepted* iterated states using the *a priori* specified proposal distribution,  $\bar{p}$ , and  $\bar{q}_\phi$ :

$$\phi^* = \arg \min_{\phi} \mathbb{E}_{p(\mathbf{x}_{t-1})} [\mathcal{D}_{\text{KL}} [\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1}) || \bar{q}_\phi(\mathbf{x}_t|\mathbf{x}_{t-1})]], \quad (2)$$

where we select the Kullback-Leibler (KL) divergence as the metric of distance between distributions. The outer expectation defines this objective as amortized across state space, where we can generate the samples by directly sampling trajectories from the model [Le et al., 2016; Gershman and Goodman, 2014]. We use the forward KL as opposed to the reverse KL,  $\mathcal{D}_{\text{KL}} [\bar{q}_\phi(\mathbf{x}_t|\mathbf{x}_{t-1}) || \bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1})]$ , as high-variance REINFORCE estimators must be used to obtain the the

differential with respect to  $\phi$  of the reverse KL.

Expanding the KL term yields:

$$\phi^* = \arg \min_{\phi} \mathbb{E}_{p(\mathbf{x}_{t-1})} \mathbb{E}_{\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1})} [\log(w)], \quad (3)$$

$$w = \frac{\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1})}{\bar{q}_\phi(\mathbf{x}_t|\mathbf{x}_{t-1})}. \quad (4)$$

Noting that  $\bar{q}_\phi$  and  $\bar{p}$  are defined only on accepted samples, where  $f$  is one-to-one, we can apply a change of variables defined for  $\bar{q}_\phi$  as:

$$\bar{q}_\phi(\mathbf{x}_t|\mathbf{x}_{t-1}) = \bar{q}_\phi(f^{-1}(\mathbf{x}_t)|\mathbf{x}_{t-1}) \left| \frac{df^{-1}(\mathbf{x}_t)}{d\mathbf{x}_t} \right|, \quad (5)$$

and likewise for  $p$ . This transforms the distribution over  $\mathbf{x}_t$  into a distribution over  $\mathbf{z}_t$  and a Jacobian term:

$$w = \frac{\bar{p}(f^{-1}(\mathbf{x}_t)|\mathbf{x}_{t-1}) \left| \frac{df^{-1}(\mathbf{x}_t)}{d\mathbf{x}_t} \right|}{\bar{q}_\phi(f^{-1}(\mathbf{x}_t)|\mathbf{x}_{t-1}) \left| \frac{df^{-1}(\mathbf{x}_t)}{d\mathbf{x}_t} \right|}. \quad (6)$$

taking care to also apply the change of variables in the distribution we are sampling from in (3). Noting that the same Jacobian terms appear in the numerator and denominator we are able to cancel these:

$$w = \frac{\bar{p}(f^{-1}(\mathbf{x}_t)|\mathbf{x}_{t-1})}{\bar{q}_\phi(f^{-1}(\mathbf{x}_t)|\mathbf{x}_{t-1})}. \quad (7)$$

We can now discard the  $\bar{p}$  term as it is independent of  $\phi$ . Noting  $f^{-1}(\mathbf{x}_t) = \mathbf{x}_{t-1} + \mathbf{z}_t$  we can write (2) as:

$$\phi^* = \arg \max_{\phi} \mathbb{E}_{p(\mathbf{x}_{t-1})} \mathbb{E}_{\bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1})} [\log \bar{q}_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})]. \quad (8)$$

However, this distribution is defined *after* rejection sampling, and can only be defined as in (1):

$$\begin{aligned} \bar{q}_\phi(\mathbf{z}_t|\mathbf{x}_{t-1}) &= q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1}, A_t = 1) \\ &= \begin{cases} \frac{1}{M_{q_\phi}} q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1}) & \text{if } f(\mathbf{x}_{t-1} + \mathbf{z}_t) \neq \perp, \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (9)$$

denoting  $M_{q_\phi}$  as the acceptance rate under  $q_\phi$ .

However, there is an infinite family of  $q_\phi$  proposals that yield  $\bar{p} = \bar{q}_\phi$ , each a maximizer of (8) but with different rejection rates. Noting however that there is only a single  $q_\phi$  that has a rejection rate of zero *and* renders  $\bar{q}_\phi = \bar{p}$ , and that this distribution also renders  $q_\phi = \bar{q}_\phi$ , we can instead optimize  $q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$ :

$$\phi^* = \arg \max_{\phi} \mathbb{E}_{p(\mathbf{x}_{t-1})} \mathbb{E}_{\bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1})} [\log q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})], \quad (10)$$

with no consideration of rejection behavior under  $q_\phi$ .

One might alternatively try to achieve low rejection rates by adding a regularization term to (8) penalizing high  $M_{q_\phi}$ . However, differentiation of  $M_{q_\phi}$  is

**Algorithm 3** Training  $q_\phi$ 


---

```

1: procedure TRAINQ( $p(\mathbf{x})$ ,  $\bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1})$ ,  $K$ ,  $N$ ,  $q$ ,
    $\phi_0$ ,  $\eta$ )
2:   for  $k = 0 : K - 1$  do
3:     for  $n = 1 : N$  do
4:        $\mathbf{x}_{t-1}^{(n)} \sim p(\mathbf{x})$   $\triangleright$  Sample from prior.
5:        $\mathbf{z}_t^{(n)} \sim \bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1}^{(n)})$   $\triangleright$  Sample noise.
6:        $E_k \leftarrow \prod_{n=1}^N q_{\phi_k}(\mathbf{z}_t^{(n)}|\mathbf{x}_{t-1}^{(n)})$ 
7:        $\mathbf{G}_k \leftarrow \nabla_{\phi_k} E_k$   $\triangleright$  Do backprop.
8:        $\phi_{k+1} \leftarrow \phi_k + \eta(\mathbf{G}_k)$   $\triangleright$  Apply update.
9:   return  $\phi_K$   $\triangleright$  Return learned parameters.
10: end procedure

```

---

intractable, meaning direct optimization of (8) is intractable.

The objective stated in (10) implicitly rewards  $q_\phi$  distributions that place minimal mass on rejections by placing as much mass on accepted samples as possible. This expression is differentiable with respect to  $\phi$  and so we can maximize this quantity through gradient ascent with minibatches drawn from  $p(\mathbf{x}_{t-1})$ . This expression shows that we can learn the distribution over accepted  $\mathbf{x}_t$  values by learning the distribution over the accepted  $\mathbf{z}_t$ , *without* needing to calculate the Jacobian or inverse of  $f$ . Doing so minimizes wasted computation, targets the same overall joint distribution, and retains interpretability by utilizing the simulator.

### 3.3 Training $q_\phi$

To train  $q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$  we first define a method for sampling state-perturbation pairs. We initialize the simulator to a state sampled from a distribution over initial value, and then iterate the perturbed simulator forward for some finite period. All state-perturbation pairs sampled during the trajectory are treated as a training example, and, in total, represent a discrete approximation to the prior over state for all time, and accepted state-conditional perturbation, i.e.  $\mathbf{x}_{t-1} \sim p(\mathbf{x})$  and  $\mathbf{z}_t \sim \bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1})$ .

We train the conditional density estimator  $q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$  using these samples by maximizing the conditional density of the sampled perturbation under the true distribution, as in (10), as this minimizes the desired KL divergence originally stated in (2). Our conditional density estimator is fully differentiable and can be trained through stochastic gradient ascent as shown in Algorithm 3. The details of the chosen architecture is explained in §3.5. The result of this procedure is an artifact that approximates the density over valid perturbations conditioned on state,  $\bar{p}(\mathbf{z}_t|\mathbf{x}_{t-1})$ .

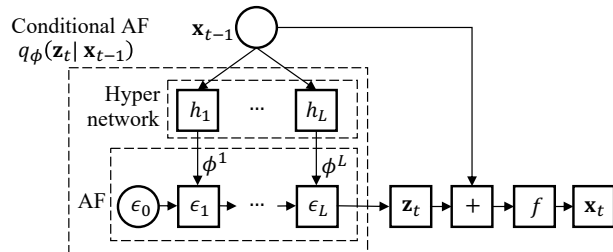


Figure 2: Diagram visualizing how  $q_\phi$  is structured and used. The previous state is input to the hypernetwork, a series of  $L$  single layer neural networks, denoted  $h_l$ . Each network outputs parameters, denoted  $\phi^l$ , for each of the  $L$  layers in the flow conditioned on the state. The flow samples a perturbation as  $\mathbf{z}_t \sim q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$ , with the internal states of the flow denoted by  $\epsilon_l$ . This perturbation is summed with the previous state and passed through the simulator,  $f$ , outputting the iterated state,  $\mathbf{x}_t$ .

### 3.4 Using $q_\phi$

Once  $q_\phi$  has been trained, it can be deployed to enhance posterior inference, by replacing samples from  $p(\mathbf{z}_t|\mathbf{x}_{t-1})$  with  $q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$ . We highlight here the ease with which it can be introduced into an SMC sweep. The state is iterated by sampling from  $\bar{p}(\mathbf{x}_t|\mathbf{x}_{t-1})$  on Line 7 of Algorithm 1, where this sampling procedure is defined in Algorithm 2. Instead of sampling from  $p(\mathbf{z}_t|\mathbf{x}_{t-1})$  in Algorithm 2, the sample is drawn from  $q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$ , and as such the sample is more likely to be accepted. This modification requires only changing a single function call made inside the implementation of Algorithm 2.

### 3.5 Implementation

We parameterize the density  $q_\phi$  using an autoregressive flow (AF) [Larochelle and Murray, 2011]. Flows define a parameterized density estimator that can be trained using stochastic gradient descent, and variants have been used in image generation [Kingma and Dhariwal, 2018], as priors for variational autoencoders [Kingma et al., 2016], and in likelihood-free inference [Papamakarios et al., 2019; Lueckmann et al., 2019].

Specifically, we structure  $q_\phi$  using a masked autoregressive flow [Papamakarios et al., 2017], with 5 single-layer MADE blocks [Germain et al., 2015], and batch normalization at the input to each intermediate MADE block. The dimensionality of the flow is the number of states perturbed in the original model. We implement conditioning through the use of a hypernetwork [Ha et al., 2016], which outputs the parameters of

the flow layers given  $\mathbf{x}_{t-1}$  as input, as shown in Figure 2. The hypernetworks are single-layer neural networks defined per flow layer. Together, the flow and hypernetwork define  $q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$ , and can be jointly trained using stochastic gradient descent. The networks are implemented in PyTorch [Paszke et al., 2017] and are optimized using ADAM [Kingma and Ba, 2014].

## 4 Experiments

### 4.1 Toy Problem – Annulus

We first demonstrate our approach on a toy problem. The true generative model of the observed data is a constant speed circular orbit around the origin in the  $x$ - $y$  plane, such that  $\mathbf{x}_t = \{x_t, y_t, \dot{x}_t, \dot{y}_t\} \in \mathbb{R}^4$ . To analyze this data we use a misspecified model that only simulates linear forward motion. To overcome the model mismatch and fit the observed data, we add Gaussian noise to position and velocity. We impose a failure constraint limiting the change in the distance of the point from the origin to a fixed threshold. This condition mirrors our observation that states in brittle simulators have large allowable perturbations in particular directions, but very narrow permissible perturbations in other directions. The true radius is unknown and so we must amortize over possible radii.

The results of this experiment are shown in Figure 3. The interior of the black dashed lines in Figure 3a indicates the permissible  $\dot{x}$ - $\dot{y}$  perturbation, for the given position and zero velocity, where we have centered each distribution on the current position for ease of visual inspection. Red contours indicate the original density  $p(\mathbf{z}_t|\mathbf{x}_{t-1})$ , and blue contours indicate the learned density  $q_\phi(\mathbf{z}_t|\mathbf{x}_{t-1})$ . The fraction of the probability mass outside the black dashed region is the expected rejection rate. Figure 3b shows the rejection rate drops from approximately 75% under the original model to approximately 4% using a trained  $q_\phi$ .

We then use the learned  $q_\phi$  as the perturbation proposal in an SMC sweep, where we condition on noisy observations of the  $x$ - $y$  coordinates. As we focus on the sample efficiency of the sweep, we fix the number of calls to the simulator in Algorithm 2 to a single call, instead of proposing and rejecting until acceptance. Failed particles are then not resampled (with certainty) during the resampling. This means that each iteration of the SMC makes a fixed number of calls to the simulator, and hence we can compare algorithms under a fixed sample budget. Figure 3c shows that we recover lower variance evidence approximations for a fixed sample budget by using  $q_\phi$  instead of  $p$ . A paired t-test evaluating the difference in variance returns a p-value of less than 0.0001, indicating a

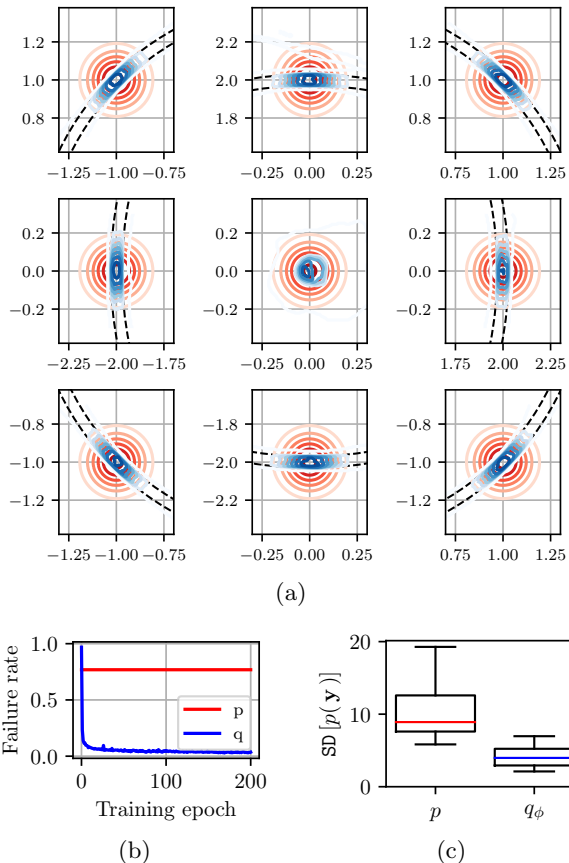


Figure 3: Results for the annulus problem introduced in Section 4.1, where the acceptable region of perturbations is inside the black dashed band. 3a shows in blue the learned state-dependent proposal distribution over velocity (for a state at rest) is the well-approximating the original proposal (shown in red) *inside* the acceptable region, with minimal spill in the invalid region, all but eliminating rejection as shown in 3b. 3c shows the reduction in the variance of the evidence by using  $q_\phi$ . We compute the variance using 100 independent SMC sweeps, each using 100 particles, and compare across 100 datasets.

strong statistical difference between the performance under  $p$  and  $q_\phi$ , confirming that using  $q_\phi$  increases the fidelity of inference for a fixed sample budget.

### 4.2 Bouncing Balls

Our second example uses a simulator of balls bouncing elastically, as shown in Figure 4a. We model the position and velocity of each ball, such that the dimensionality of the state vector,  $\mathbf{x}_t$ , is four times the number of balls. We add a small amount of Gaussian noise at each iteration to the position and velocity of each ball. This perturbation induces the possibility that two balls overlap, or, a ball intersects with the wall,

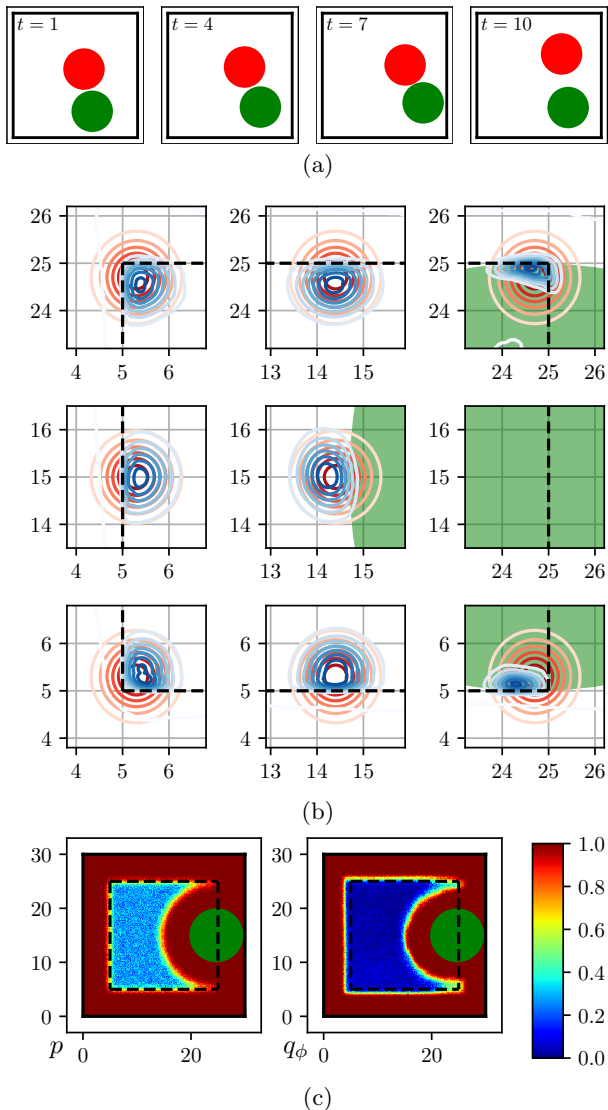


Figure 4: Results of the bouncing balls experiment introduced in Section 4.2, with two radius five, unit mass balls in an enclosure of size 30. 4a shows an example trajectory of the system. 4b shows, in red, the proposal distribution over the perturbation to the position of the first ball specified in the model, and the learned proposal in blue. The edge of the permissible region of the enclosure is shown as a black dashed line. The second ball is fixed at  $[25, 15]$ , and the induced invalid region shaded green. The flow has learned to deflect away from the disallowed regions. 4c shows the rejection rate as a function of the position of the first ball, with the second ball in the position shown. The trained proposal (right) has all but eliminated rejection in the permissible space compared to the a-priori specified proposal (left). The rejection rate under  $p$  is high in the interior as the second ball may also leave the enclosure, whereas  $q_\phi$  has practically eliminated rejection by *jointly* proposing perturbations.

representing an invalid physical configuration and results in simulator failure. We note that here, we are conditioning on the state of *all* balls simultaneously, and proposing the perturbation to the state *jointly*.

Figure 4b shows the distribution over position perturbation of a single ball, conditioned on the other ball being stationary. Blue contours show the estimated distribution over accepted perturbations learned by autoregressive flow. Figure 4c shows the rejection rate under  $p$  and  $q_\phi$  as a function of the position of the first ball, with the second ball fixed in the position shown, showing that rejection has been all but eliminated. We again see a reduction in the variance of the evidence approximation computed by a particle filter when using  $q_\phi$  instead of  $p$  (figure in the supplementary materials).

### 4.3 MuJoCo

We now apply our method to the popular robotics simulator MuJoCo [Todorov et al., 2012], specifically using the built-in example “tossler,” where a capsule is “tossed” by an actuator into a bucket, shown in Figure 5a. Tossler displays “chaotic” aspects, as minor changes in the position of the object results in large changes in the trajectories achieved by the simulator.

MuJoCo allows some overlap between the objects to simulate contact dynamics. This is an example of model misspecification borne out of the requirements of reasonably writing a simulator. We therefore place a hard limit on the amount objects are allowed to overlap. This is an example of a user-specified constraint that requires the simulator to be run to evaluate. We add Gaussian distributed noise to the position and velocity of the capsule.

Figure 5 shows the results of this experiment. The capsule is mostly in free space resulting in an average rejection rate under  $p$  of 10%. Figure 5b shows that the autoregressive flow learns a proposal with a lower rejection rate, reaching 3% rejection. However these rejections are concentrated in the critical regions of state-space, where chaotic behavior occurs, and so this reduction yields an large reduction in the variance of the evidence approximation, as shown in Figure 5c.

We conclude this example by evaluating our method on hypothesis testing using pseudo-marginal evidence estimates. The results for this are shown in Figure 5d. We test 5 different hypothesis of the mass of the capsule. Using  $p$  results in higher variance evidence approximations than when  $q_\phi$  is used. Additionally, under  $p$  the wrong model is selected (2 instead of 3), although with low significance ( $p = 0.125$ ), while using  $q_\phi$  selects the correct hypothesis with  $p = 0.0127$ . For this experiment we note that  $q_\phi$  was trained on

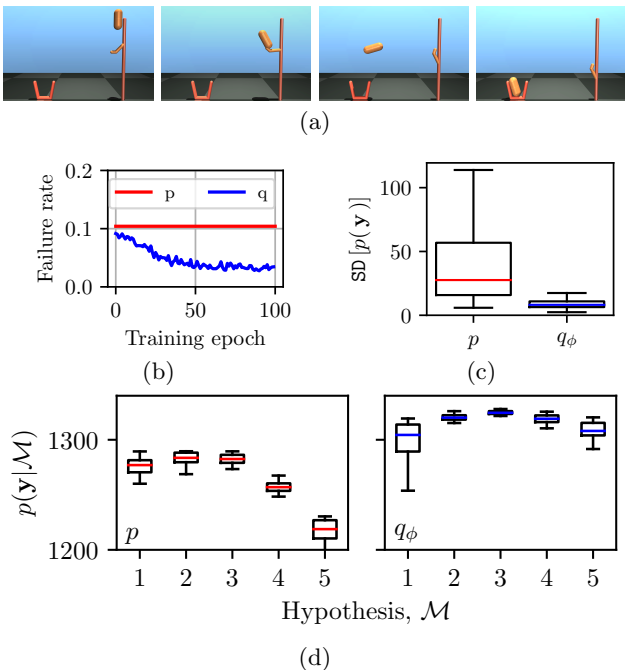


Figure 5: Results of the “tossers” experiment introduced in Section 4.3. 5a shows the evolution of state over time. 5b shows the AF we learn markedly reduces the number of rejections. 5c shows the results of performing SMC using the *a priori* specified proposal and our learned autoregressive flow. The autoregressive flow attains a much lower variance estimate, with a p-value of less than 0.0001 in a paired t-test, indicating a strong statistical difference in performance. 5d shows the results of performing hypothesis testing, where hypothesis 3 is correct, under a uniform prior over hypothesis. The incorrect hypothesis is selected using  $p$ , while using  $q_\phi$  the correct hypothesis is selected, with a statistically significant confidence.

a single value of mass, and that this “training mass” was different to the “testing mass.” We believe this contributes to the increased variance in hypothesis 1, which is very light compared to the training mass. Training a  $q_\phi$  with a further level of amortization over different mass values would further increase the fidelity of the model selection. This is intimately linked with the larger project of jointly learning the model, and so we defer investigation to future works.

#### 4.4 Neuroscience Simulator

We conclude by applying our algorithm to a simulator for the widely studied *Caenorhabditis elegans* roundworm. WormSim, presented by Boyle et al. [2012], is a simulator of the locomotion of the worm, using a 510 dimensional state representation. We apply perturbations to a 98 dimensional subspace defining the physical position of the worm, while conditioning on

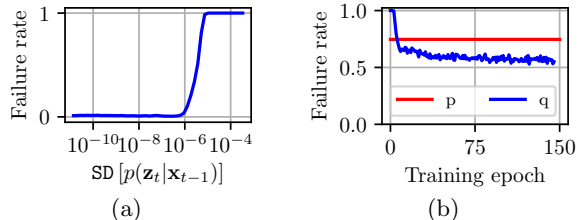


Figure 6: Results from the WormSim example introduced in Section 4.4. 6a shows the rate at which the simulator fails increases sharply as a function of the standard deviation of the applied perturbation. 6b shows the reduction in rejections during training.

the full 510 dimensional state vector. The expected rate of failure increases sharply as a function of the scale of the perturbation applied, as shown in Figure 6a, as the integrator used in WormSim is unable to integrate highly perturbed states.

The rejection rate during training is shown in Figure 6b. We are able to learn an autoregressive flow with lower rejection rates, reaching approximately 53% rejection, when  $p$  has approximately 75% rejection. Although the rejection rate is higher than ultimately desired, we include this example as a demonstration of how rejections occur in simulators through integrator failure. We believe larger flows with regularized parameters can reduce the rejection rate further.

## 5 Conclusion

In this paper we have tackled reducing simulator failures caused by naively perturbing the input state. We achieve this by showing that stochastically perturbed simulators define a rejection sampler with a well defined target distribution and learning a conditional autoregressive flow to estimate the state-dependent proposal distribution conditioned on acceptance. We then show that using this learned proposal reduces the variance of inference results, with applications for Bayesian model selection. We believe this work has readily transferable practical contributions to not just the machine learning community, but the wider scientific community where such naively modified simulation platforms are widely deployed. As part of the experiments we present, we identify an extension: introducing an additional level of amortization over static simulation parameters. This extension builds towards our larger research vision of building toolchains for efficient inference and learning in brittle simulators. Further development will facilitate efficient gradient-based model learning in these brittle simulators.



## 6 Acknowledgements

Andrew Warrington is supported under the Shilston Scholarship, awarded by Keble College and the Department of Engineering Science, University of Oxford. Saeid Naderiparizi and Frank Wood are supported under the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada CIFAR AI Chairs Program, Compute Canada, Intel, Composites Research Network (CRN) and DARPA under its D3M and LWLL programs.

## References

- R. Adhikari and R. K. Agrawal. An introductory study on time series modeling and forecasting. *arXiv preprint arXiv:1302.6613*, 2013.
- L. J. Allen. A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infectious Disease Modelling*, 2(2):128–142, 2017.
- J. H. Boyle, S. Berri, and N. Cohen. Gait modulation in *c. elegans*: an integrated neuromechanical model. *Frontiers in computational neuroscience*, 6:10, 2012.
- P. J. Brockwell and R. A. Davis. *Introduction to time series and forecasting*. springer, 2016.
- L. Coutin, J.-M. Guglielmi, and N. Marie. On a fractional stochastic hodgkin–huxley model. *International Journal of Biomathematics*, 11(05):1850061, 2018.
- A. Doucet, N. De Freitas, and N. Gordon. An introduction to sequential monte carlo methods. In *Sequential Monte Carlo methods in practice*, pages 3–14. Springer, 2001.
- A. Doucet, S. J. Godsill, and C. P. Robert. Marginal maximum a posteriori estimation using markov chain monte carlo. *Statistics and Computing*, 12(1):77–84, 2002.
- N. H. Du and V. H. Sam. Dynamics of a stochastic lotka–volterra model perturbed by white noise. *Journal of mathematical analysis and applications*, 324(1):82–97, 2006.
- N. R. Edwards, D. Cameron, and J. Rougier. Precalibrating an intermediate complexity climate model. *Climate dynamics*, 37(7-8):1469–1482, 2011.
- M. F. Fallon, H. Johannsson, and J. J. Leonard. Efficient scene simulation for robust monte carlo localization using an rgb-d camera. In *2012 IEEE international conference on robotics and automation*, pages 1663–1670. IEEE, 2012.
- R. F. Fox. Stochastic versions of the hodgkin-huxley equations. *Biophysical journal*, 72(5):2068–2074, 1997.
- M. Germain, K. Gregor, I. Murray, and H. Larochelle. Made: Masked autoencoder for distribution estimation. In *International Conference on Machine Learning*, pages 881–889, 2015.
- S. Gershman and N. Goodman. Amortized inference in probabilistic reasoning. In *Proceedings of the annual meeting of the cognitive science society*, volume 36, 2014.
- J. H. Goldwyn and E. Shea-Brown. The what and where of adding channel noise to the hodgkin-huxley equations. *PLOS Computational Biology*, 7(11):1–9, 11 2011.
- D. Ha, A. Dai, and Q. V. Le. Hypernetworks. *arXiv preprint arXiv:1609.09106*, 2016.
- R. E. Kalman et al. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.
- N. Kantas, A. Doucet, S. S. Singh, J. Maciejowski, N. Chopin, et al. On particle methods for parameter estimation in state-space models. *Statistical science*, 30(3):328–351, 2015.
- D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- D. P. Kingma and P. Dhariwal. Glow: Generative flow with invertible 1x1 convolutions. In *Advances in Neural Information Processing Systems*, pages 10215–10224, 2018.
- D. P. Kingma, T. Salimans, R. Jozefowicz, X. Chen, I. Sutskever, and M. Welling. Improved variational inference with inverse autoregressive flow. In *Advances in neural information processing systems*, pages 4743–4751, 2016.
- H. Larochelle and I. Murray. The neural autoregressive distribution estimator. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 29–37, 2011.
- T. A. Le, A. G. Baydin, and F. Wood. Inference compilation and universal probabilistic programming. *arXiv preprint arXiv:1610.09900*, 2016.
- H. F. Lopes and R. S. Tsay. Particle filters and bayesian inference in financial econometrics. *Journal of Forecasting*, 30(1):168–209, 2011. doi: 10.1002/for.1195.
- D. Lucas, R. Klein, J. Tannahill, D. Ivanova, S. Brandon, D. Domyancic, and Y. Zhang. Failure analysis of parameter-induced simulation crashes in climate models. *Geoscientific Model Development*, 6(4):1157–1171, 2013.
- J.-M. Lueckmann, G. Bassetto, T. Karaletsos, and J. H. Macke. Likelihood-free inference with emu-

- lator networks. In *Symposium on Advances in Approximate Bayesian Inference*, pages 32–53, 2019.
- Q. Lv, M. K. Schneider, and J. W. Pitchford. Individualism in plant populations: Using stochastic differential equations to model individual neighbourhood-dependent plant growth. *Theoretical Population Biology*, 74(1):74 – 83, 2008. ISSN 0040-5809. doi: <https://doi.org/10.1016/j.tpb.2008.05.003>.
- D. J. MacKay. *Information theory, inference and learning algorithms*. Cambridge university press, 2003.
- I. S. Mbalawata, S. Särkkä, and H. Haario. Parameter estimation in stochastic differential equations with markov chain monte carlo and non-linear kalman filtering. *Computational Statistics*, 28(3):1195–1223, Jun 2013. ISSN 1613-9658. doi: [10.1007/s00180-012-0352-y](https://doi.org/10.1007/s00180-012-0352-y).
- J. K. Møller, H. Madsen, and J. Carstensen. Parameter estimation in a simple stochastic differential equation for phytoplankton modelling. *Ecological Modelling*, 222(11):1793–1799, 2011.
- G. Papamakarios, T. Pavlakou, and I. Murray. Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems*, pages 2338–2347, 2017.
- G. Papamakarios, D. Sterratt, and I. Murray. Sequential neural likelihood: Fast likelihood-free inference with autoregressive flows. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 837–848, 2019.
- A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in PyTorch. In *NIPS Autodiff Workshop*, 2017.
- S. M. Pimblott and J. A. LaVerne. Comparison of stochastic and deterministic methods for modeling spur kinetics. *Radiation Research*, 122(1):12–23, 1990. ISSN 00337587, 19385404.
- S. Razavi, R. Sheikholeslami, H. V. Gupta, and A. Haghnegahdar. Vars-tool: A toolbox for comprehensive, efficient, and robust sensitivity and uncertainty analysis. *Environmental Modelling & Software*, 112:95 – 107, 2019. ISSN 1364-8152.
- K. Reddy and V. Clinton. Simulating stock prices using geometric brownian motion: Evidence from australian companies. *Australasian Accounting, Business and Finance Journal*, 10(3):23–47, 2016.
- P. Renard, A. Alcolea, and D. Gingsbourger. Stochastic versus deterministic approaches. In *Environmental Modelling: Finding Simplicity in Complexity, Second Edition (eds J. Wainwright and M. Mulligan)*, pages 133–149. Wiley Online Library, 2013.
- A. Saarinen, M.-L. Linne, and O. Yli-Harja. Stochastic differential equation model for cerebellar granule cell excitability. *PLOS Computational Biology*, 4(2):1–11, 02 2008. doi: [10.1371/journal.pcbi.1000004](https://doi.org/10.1371/journal.pcbi.1000004).
- R. Sheikholeslami, S. Razavi, and A. Haghnegahdar. What do we do with model simulation crashes? recommendations for global sensitivity analysis of earth and environmental systems models. *Geoscientific Model Development Discussions*, 2019:1–32, 2019. doi: [10.5194/gmd-2019-17](https://doi.org/10.5194/gmd-2019-17).
- S. Thrun, D. Fox, W. Burgard, and F. Dellaert. Robust monte carlo localization for mobile robots. *Artificial intelligence*, 128(1-2):99–141, 2001.
- E. Todorov, T. Erez, and Y. Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012.