

Quantifying mobile-ion parameters in perovskite solar cells using tempered sequential Monte Carlo and an evidential neural network surrogate

Lewis Dunning

Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom

E-mail: ld819@bath.ac.uk

Abstract. Perovskite solar cells (PSCs) achieve high power conversion efficiencies but remain limited by operational instability, in which mobile ionic defects redistribute under bias and illumination. In metal-halide PSCs, mobile halide vacancies drift and accumulate at transport-layer interfaces, forming interfacial Debye layers that screen the internal electric field, drive scan-rate-dependent J - V hysteresis, and can reduce settled (steady-state) performance during ageing. Physics-based drift-diffusion models capture this coupled electronic-ionic behaviour, but key ionic parameters – the mobile vacancy density N_0 and diffusion coefficient D_I – are difficult to measure directly and are typically inferred by reverse modelling from electrical data. Here, we infer N_0 and D_I (together with selected electronic and transport-layer parameters) from multiple scan-rate J - V measurements using Bayesian parameter estimation with a tempered Sequential Monte Carlo (SMC) sampler featuring adaptive tempering and proposal adaptation. To enable rapid, uncertainty-aware inference, we also train an evidential neural network (ENN) surrogate on synthetic IonMonger simulations to provide single-pass parameter estimates with predictive uncertainty. Across an illumination-ageing J - V dataset, SMC yields the lowest reconstruction error and recovers a clear ionic ageing signature: N_0 increases from $\sim 7 \times 10^{21}$ to $\sim 2 \times 10^{25} \text{ m}^{-3}$ while D_I decreases from $\sim 6 \times 10^{-13}$ to $\sim 5 \times 10^{-16} \text{ m}^2 \text{ s}^{-1}$ over 480 min, consistent with strengthened ion-induced field screening. For impedance spectroscopy, we test on a synthetic spectrum generated by IonMonger at a fixed DC operating point (small-signal linearisation) using the same fixed device constants as those assumed during ENN training; under this matched setting, the ENN reproduces the target spectrum most closely at a fraction of the computational cost. Overall, the ENN is effective for rapid screening and uncertainty flagging, while targeted SMC inference remains preferable when model mismatch or strongly correlated posteriors require a fully Bayesian treatment.

1. Introduction

The global transition away from fossil fuels is accelerating renewable generation, with solar power expanding rapidly due to scalability and declining cost. [1, 2] Crystalline silicon remains the dominant photovoltaic technology and a key performance benchmark. [3–5] In parallel, metal-halide perovskites have emerged as outstanding optoelectronic semiconductors, enabling rapid efficiency gains from 3.8% in 2009 to ~ 26 –27% certified single-junction power conversion efficiency (PCE) and record tandem architectures. [3, 5–11]

Beyond high efficiency, perovskites combine strong absorption and tunable bandgaps with low-temperature processing, supporting thin-film, tandem, semi-transparent and flexible designs. [7, 8, 11–13] However, operational stability remains a central barrier to deployment. [11, 14, 15] Among coupled degradation processes, mobile ionic defects are particularly important because they redistribute under bias/illumination, screen internal fields, and drive scan-rate-dependent J - V hysteresis and associated performance losses. [16–20]

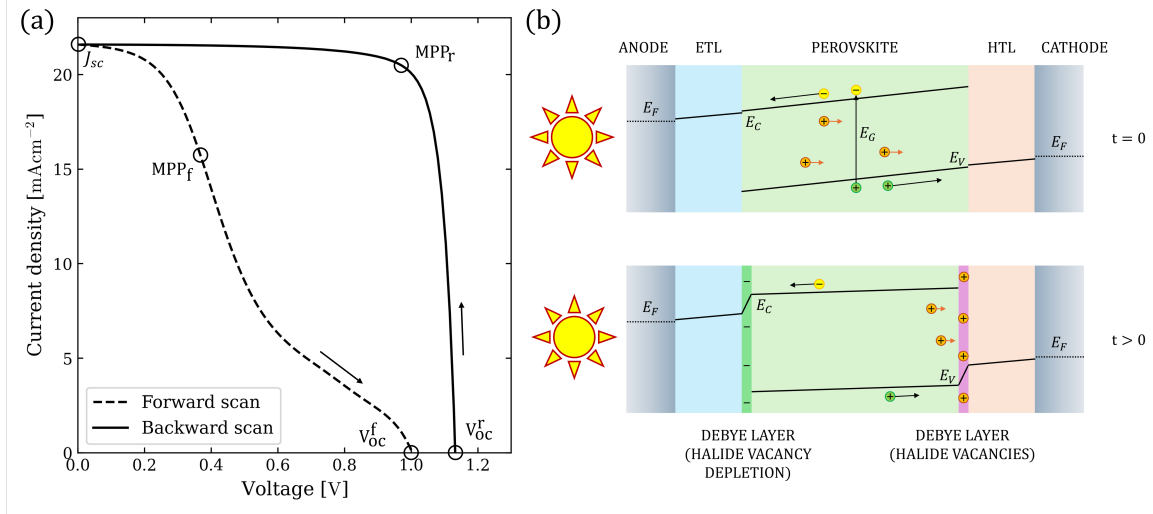


Figure 1. Scan-rate-dependent J - V hysteresis in perovskite solar cells and its ionic origin. (a) Reverse (solid) and forward (dashed) scans show similar J_{sc} but different apparent V_{oc} and MPP due to measurement-history-dependent internal fields. (b) Positively charged halide vacancies drift under the field and accumulate at the perovskite/ETL and perovskite/HTL interfaces, forming Debye layers that screen the bulk field and redistribute the potential drop. The resulting changes in band bending, extraction and recombination differ between scan directions, producing hysteresis.

A key distinction between perovskites and conventional inorganic semiconductors is the presence of halide vacancies, of which there is a large concentration due to low formation energy. The perovskite layer behaves as a weak solid electrolyte, and halide-related species migrate under internal and applied fields. [16–20] Drift-diffusion modelling and electrical measurements show that ion vacancies and charged defects accumulate at transport-layer interfaces, forming narrow Debye layers that partially screen the internal field. [16, 21, 22] Cave *et al.* demonstrated that including ion vacancy motion and both transport layers reproduces the scan-rate and temperature dependence of J - V curves, and that hysteresis arises mainly from halide ions entering and leaving these Debye layers. [16, 19, 21, 23] They further showed that such screening can depress settled PCE even when hysteresis is small, and that lowering ion vacancy density improves long-term performance. [21, 24]

Mobile ions affect not only hysteresis but also degradation. Thiesbrummel *et al.* combined accelerated ageing with electrical characterisation, exposing PSCs to controlled illumination, heat and bias while monitoring J - V curves and capacitance. [20] Drift-diffusion fits to the transients indicated that mobile ion vacancy densities can increase by several orders of magnitude under light-induced ageing, and that this alone can explain the observed drop in steady-state PCE through enhanced ionic screening of the internal field. [16, 19, 20] They introduced the notion of “ionic losses”—PCE reductions specifically due to ion-induced field screening—which can amount to several percentage points even in efficient devices. [20] Together with earlier studies of ion migration and interfacial reactions, this underscores the need for quantitative characterisation of mobile ion populations to understand and mitigate PSC degradation. [14, 16, 17, 20]

The impact of mobile ions is also clear in detailed device simulations. Open-source drift-diffusion tools such as IONMONGER explicitly model coupled electron, hole and ion vacancy transport in planar PSCs. [25, 26] Using such models, Clarke *et al.* studied degradation, represented as an increasing recombination rate, and its effect on J - V curves and impedance spectra. [27] They showed that impedance spectroscopy can resolve PCE losses as small as $\sim 0.1\%$ over hours and that Nyquist loops may arise from either ionic dynamics in pristine cells or rapid degradation during measurement. [23, 27] However, these physics-based models rely on material parameters—particularly ion vacancy density and diffusion coefficient—that are hard to measure directly.

To address this, McCallum *et al.* proposed a Bayesian parameter estimation (BPE) framework

for characterising mobile ion vacancies. [22] IONMONGER is used to simulate multi-scan-rate current-voltage, J - V , curves for given ion vacancy density N_0 and diffusion coefficient D_I , and a Metropolis–Hastings MCMC sampler infers their posterior distributions from experimental data. [22] This reverse-modelling approach yields quantitative estimates of N_0 and D_I with credible intervals and recovers parameters to within roughly an order of magnitude over a broad range. [22] Its main limitation is computational cost: multiple long MCMC chains are required, and efficiency is sensitive to the proposal distribution, as is typical for Metropolis–Hastings algorithms. [22, 28]

In this work, we use IONMONGER to infer key device parameters from measured J - V and impedance data. We apply Bayesian parameter estimation using tempered Sequential Monte Carlo and develop an evidential neural network to provide fast, uncertainty-aware parameter estimates. We test these approaches on an illumination ageing J - V dataset and on impedance spectroscopy data, and compare their accuracy and computational cost.

2. Theory

2.1. Forward model

Consider a planar perovskite solar cell consisting of an electron-transport layer (ETL), an absorber layer (the perovskite) and a hole-transport layer (HTL), arranged in a one-dimensional stack. [25] The perovskite is treated as a mixed ionic–electronic conductor: electrons and holes move by drift and diffusion, while a single species of mobile ion vacancies (e.g. iodide vacancies) migrates under the electric field and concentration gradients. [25]

The model consists of coupled electron and hole continuity (drift–diffusion) equations and an ionic Nernst–Planck equation, closed by Poisson’s equation for the electrostatic potential. [25] Photogeneration in the perovskite is described by a Beer–Lambert law, with a constant absorption coefficient and a spatially varying generation rate that decays exponentially away from the illuminated interface. [25] Bulk recombination is represented by a Shockley–Read–Hall (SRH) mechanism with a single deep trap, using material parameters and functional form as in Courtier *et al.* [25]

At the ETL/perovskite and perovskite/HTL interfaces, continuity of current and appropriate band offsets are imposed, together with surface recombination velocities to account for interfacial recombination losses. [25] Fixed ionic charge in the transport layers and mobile vacancies in the perovskite give rise to narrow Debye layers in which most of the potential drop occurs, a key ingredient for reproducing hysteretic J - V behaviour. [25]

To mimic quasi-static J - V measurements, the device is first preconditioned at $V_{\text{pre}} = 1.2$ V under illumination for 40 s to allow ionic charge to relax near the contacts. From this state, a reverse scan from 1.2 V to 0 V is applied at a fixed rate, followed immediately by a forward scan back to 1.2 V; repeating this at different scan rates generates a family of hysteretic J - V curves. [25] Rather than comparing simulated and measured J - V curves pointwise, we summarise each pair of forward and reverse scans at a given scan rate using:

- J_{sc} : short-circuit current density, i.e. at $V = 0$ on the reverse scan;
- V_{oc}^{r} and V_{oc}^{f} : open-circuit voltages on reverse and forward scans;
- MPP_{r} and MPP_{f} : maximum-power-points on reverse and forward scans.

The same forward model is used to compute small-signal impedance spectra by applying a sinusoidal voltage perturbation about a chosen operating point and recording the complex impedance $Z(f_k)$ at 64 frequencies. These J - V and impedance summaries form the input to the SMC algorithm.

2.2. Statistical Model

We collect all unknown model parameters in a vector $\boldsymbol{\theta}$, and assign each component independent uniform priors on physically reasonable bounds $[a_j, b_j]$. For numerical stability we map to an unconstrained internal vector z by a componentwise logit transform of the rescaled θ_j , so that each

Hyperparameter	Value
Particle number N	800
Log-likelihood denominator σ^2	1×10^{-5}
Effective sample size threshold ESS	0.5
Random Walk Metropolis–Hastings steps	4
Conditional effective sample size target ϕ	0.99
Proposal step scale γ_t	1.5
Target acceptance rate p^*	0.28

Table 1. Chosen hyperparameters for the SMC sampler.

bounded $\theta_j \in [a_j, b_j]$ corresponds to $z_j \in \mathbb{R}$. The induced prior density on z is added (via its log-density) to the log-likelihood to form the log-posterior used by the SMC sampler.

For either J – V or impedance data we work with summary vectors $s(\theta)$ (model) and s_{obs} (data), and measure misfit using the mean-squared error (MSE) between these summaries. Assuming independent Gaussian observation errors with common variance, the log-likelihood is proportional to minus the MSE, with a single variance scale σ^2 controlling the overall weight given to the data.

2.3. Sequential Monte Carlo tempering

We introduce a tempered sequence of target distributions $\{\pi_\lambda(\theta)\}_{\lambda \in [0,1]}$ of the form

$$\pi_\lambda(\theta) \propto p(\theta) \exp(\lambda \ell(\theta)), \quad 0 = \lambda_0 < \lambda_1 < \dots < \lambda_T = 1,$$

where $p(\theta)$ is the prior and $\ell(\theta)$ the log-likelihood defined above. At $\lambda_0 = 0$ the target is the prior, and at $\lambda_T = 1$ it is the posterior. SMC samplers approximate this sequence by a weighted particle system $\{(\theta_t^{(i)}, w_t^{(i)})\}_{i=1}^N$ at each level λ_t [29, 30].

Given a sample approximating $\pi_{\lambda_{t-1}}$, we move to $\lambda_t = \lambda_{t-1} + \Delta\lambda_t$ in three standard SMC steps: reweighting, resampling (if needed), and mutation. For reweighting, we evaluate $\ell(\theta_{t-1}^{(i)})$ (if not already stored) and compute

$$v_t^{(i)} = \exp(\Delta\lambda_t \ell(\theta_{t-1}^{(i)})), \quad \tilde{w}_t^{(i)} = w_{t-1}^{(i)} v_t^{(i)}, \quad w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^N \tilde{w}_t^{(j)}}.$$

Weight degeneracy is monitored using the effective sample size (ESS) [31]. When the ESS falls below a fraction $\alpha_{\text{ESS}}N$ (typically $\alpha_{\text{ESS}} \approx 0.5$), we resample using systematic resampling [30], after which $w_t^{(i)} = 1/N$. To rejuvenate the particle cloud and maintain good exploration of π_{λ_t} , we then apply several random walk Metropolis–Hastings (RWMH) updates that leave π_{λ_t} invariant, using a Gaussian proposal in the internal (z -) parameterisation.

Choosing the tempering increments $\Delta\lambda_t$ by hand is difficult and problem-specific, so we adapt them using the conditional effective sample size (CESS) criterion [31, 32]. For a candidate increment $\Delta\lambda_t$ we define incremental weights

$$v^{(i)}(\Delta\lambda_t) = \exp(\Delta\lambda_t \ell(\theta_{t-1}^{(i)})),$$

and compute

$$\text{CESS}(\Delta\lambda_t) = N \frac{(\sum_{i=1}^N w_{t-1}^{(i)} v^{(i)}(\Delta\lambda_t))^2}{\sum_{i=1}^N w_{t-1}^{(i)} (v^{(i)}(\Delta\lambda_t))^2}.$$

At each iteration we choose $\Delta\lambda_t$ (e.g. by bisection) so that $\text{CESS}(\Delta\lambda_t) = \phi N$ for a user-specified fraction $\phi \in (0, 1)$ (typically $0.8 \leq \phi \leq 0.99$), giving an automatically adapted tempering schedule with controlled weight degeneracy [31, 32].

The RWMH proposal covariance at level t is written

$$\Sigma_t = s_t^2 \Lambda_t,$$

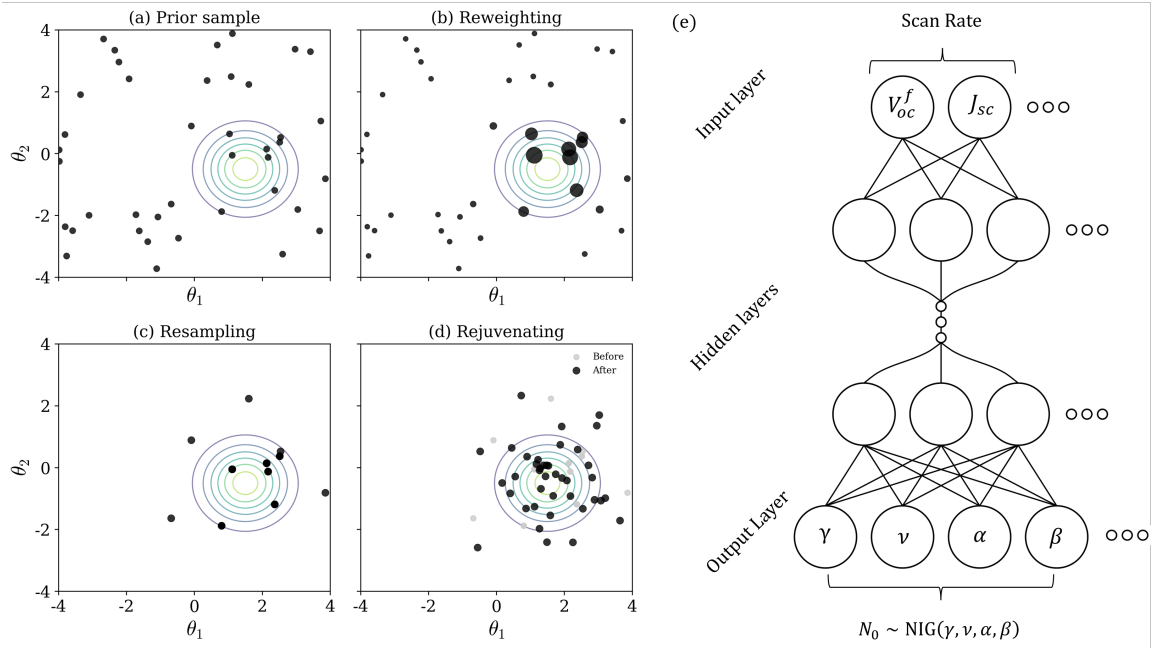


Figure 2. Tempered SMC overview and ENN schematic. (a–d) Particles are reweighted, resampled, and rejuvenated with MH moves along a tempering schedule from prior to posterior. (e) ENN maps simulated features to parameter estimates with uncertainty via NIG outputs. The circles represent nodes and the lines, weights.

where s_t is a scalar step size and Λ_t a symmetric positive-definite shape matrix capturing anisotropy and correlations [33,34]. After each batch of MH moves we compute the empirical acceptance rate \hat{p}_t and update the scale via a Robbins–Monro recursion [33,35]:

$$\log s_{t+1} = \log s_t + \gamma_t (\hat{p}_t - p^*),$$

where γ_t is a decreasing gain sequence and p^* a target acceptance rate (e.g. $p^* \approx 0.2\text{--}0.3$). After each resampling step we estimate Λ_t from the current weighted particle cloud in the internal parameterisation and apply shrinkage towards a fixed diagonal base covariance, which stabilises the estimate when N is modest or the posterior is strongly correlated, while still allowing the proposal to adapt to the evolving target [34].

2.4. Evidential neural network surrogate (ENN)

As an alternative to statistical approaches such as the Sequential Monte Carlo (SMC) sampler, we use a deterministic neural-network surrogate that maps device characteristics directly to underlying physical parameters while also estimating predictive uncertainty. This surrogate is based on evidential deep learning (EDL), a single-forward-pass framework for uncertainty quantification in deep neural networks [36]. EDL interprets network outputs as *evidence* supporting different parameter values via subjective logic and evidence theory [36–38]. Compared with Bayesian neural networks or deep ensembles, which require multiple stochastic forward passes, EDL provides fast, deterministic uncertainty estimates.

The evidential neural network (ENN) takes as input a feature vector derived from IONMONGER simulations and outputs point estimates of device parameters together with associated uncertainties. To generate training data, we convert the prior ranges in table 2 to \log_{10} space, draw parameter vectors uniformly in this space, and pass them to IONMONGER to simulate J – V or IS responses. For J – V , five scan rates (0.25, 1.25, 25, 250 and 2500 V s^{-1}) are used, from which J_{sc} , V_{oc}^r , V_{oc}^f , MPP_r and MPP_f are extracted. For IS, the input features are resistance and reactance at 64 frequencies between 0.1 mHz and 1 MHz, uniformly spaced in log-space. In total, 100 000 J – V samples and 5 000

Parameter θ	J-V range	IS range	Unit	Ref.
Mobile ion vacancy density N_0	10^{20} – 10^{26}	10^{24} – 10^{26}	m^{-3}	[22]
Mobile ion diffusion coefficient D_I	10^{-16} – 10^{-11}	10^{-18} – 10^{-15}	$\text{m}^2 \text{s}^{-1}$	
ETL doping density d_E	10^{19} – 10^{26}	–	m^{-3}	
HTL doping density d_H	10^{20} – 10^{25}	10^{23} – 10^{25}	m^{-3}	
SRH electron recombination velocity at ETL v_n^E	10^{-1} – 10^5	–	m s^{-1}	
SRH hole recombination velocity at ETL v_p^E	10^{-3} – 10^3	–	m s^{-1}	
SRH electron recombination velocity at HTL v_n^H	10^{-3} – 10^3	–	m s^{-1}	
SRH hole recombination velocity at HTL v_p^H	10^{-1} – 10^5	–	m s^{-1}	
Electron diffusion coefficient D_n	–	10^{-4} – 10^{-2}	$\text{m}^2 \text{s}^{-1}$	
Hole diffusion coefficient D_p	–	10^{-4} – 10^{-2}	$\text{m}^2 \text{s}^{-1}$	
ETL permittivity ε_E	–	10^{-11} – 10^{-10}	Fm^{-1}	
HTL permittivity ε_H	–	10^{-11} – 10^{-10}	Fm^{-1}	

Table 2. Prior parameter ranges used to generate synthetic J-V and IS datasets.

IS samples are generated, the smaller IS set reflecting the higher IONMONGER runtime (about $20\times$ that of J - V). As accurate training data are essential, nonphysical outputs are removed: J - V cases with negative open-circuit voltages are discarded, and IS spectra with discontinuous gradients or diverging impedances are filtered. An automated filter removes about 95% of clearly unphysical IS results; the remaining problematic cases are removed manually.

Standard neural networks are typically trained with a mean-squared-error (MSE) loss, which does not provide uncertainty estimates. In contrast, ENNs use an evidence-based regularisation and learn the parameters of normal-inverse-gamma (NIG) distributions, from which both predictions and uncertainties can be derived. For each scalar target the ENN outputs

$$m = (\gamma, \nu, \alpha, \beta),$$

which parameterise a NIG distribution over the Gaussian likelihood parameters (μ, σ^2) [36, 39]. The ENN loss function is

$$\mathcal{L}(w) = \mathcal{L}_{\text{NLL}}(w) + \lambda \mathcal{L}_R(w), \quad (1)$$

where w is the vector of network weights and λ is a regularisation prefactor. The term \mathcal{L}_{NLL} encourages the NIG parameters to fit the data, while \mathcal{L}_R penalises precise (low-uncertainty) but incorrect predictions.

Introducing $\Omega = 2\beta(1 + \nu)$, and denoting the target by y_i , the components of the loss are

$$\mathcal{L}_{\text{NLL}} = \frac{1}{2} \log\left(\frac{\pi}{\nu}\right) - \alpha \log(\Omega) + \left(\alpha + \frac{1}{2}\right) \log[(y_i - \gamma)^2 \nu + \Omega] + \log\left(\frac{\Gamma(\alpha)}{\Gamma(\alpha + \frac{1}{2})}\right), \quad (2)$$

$$\mathcal{L}_R = |y_i - \gamma| (2\nu + \alpha), \quad (3)$$

where Γ is the Gamma function. From the NIG distribution, the mean, aleatoric uncertainty and epistemic uncertainty of the underlying normal likelihood are

$$\text{E}(\mu) = \gamma, \quad \text{E}(\sigma^2) = \frac{\beta}{\alpha - 1}, \quad \text{Var}(\mu) = \frac{\beta}{\nu(\alpha - 1)}. \quad (4)$$

Here μ and σ^2 are the mean and variance of the Gaussian likelihood; E denotes expectation and Var variance. The epistemic (model) uncertainty $\text{Var}(\mu)$ reflects lack of knowledge and can be reduced by more data or a better model, whereas the aleatoric (data) uncertainty $\text{E}(\sigma^2)$ reflects irreducible noise [40].

Because each device parameter requires one NIG distribution, and each NIG is specified by four parameters, the ENN output layer is four times larger than that of an otherwise comparable deterministic network. Hyperparameters controlling the ENN architecture and training (hidden

Hyperparameter	Range	J-V model	IS model
Hidden dimension	64–1024	768	512
Hidden layers	4–17	8	7
Dropout	0.05–0.5	0.28	0.19
Learning rate	10^{-5} – 10^{-2}	5.7×10^{-3}	1.5×10^{-4}
λ	10^{-4} – 10^{-1}	1.8×10^{-3}	2.1×10^{-3}
Weight decay	10^{-6} – 10^{-3}	6.4×10^{-5}	8.1×10^{-5}

Table 3. Hyperparameters selected for the J-V and IS ENN models.

dimension, number of layers, dropout, learning rate, λ and weight decay) are tuned using Optuna with 250 trials over specified ranges.

Training uses the AdamW optimiser with weight decay and a cosine-annealing learning-rate schedule, which helps the model escape poor local minima. Gradient clipping is applied for stability. The regularisation weight is annealed during training [36] according to

$$\lambda_e = \lambda_{\text{target}} \min\left(1, \frac{e}{E_{\text{anneal}}}\right),$$

where e is the epoch index, λ_{target} is the target regularisation strength and E_{anneal} is the number of epochs over which λ_e ramps from 0 to λ_{target} . Training is continued until the validation loss ceases to improve (early stopping), indicating the onset of overfitting.

3. Results and Discussion

3.1. Bayesian parameter estimation of the ageing series

Figure 3 shows posterior-peak parameters inferred using SMC Bayesian parameter estimation (BPE) applied to the Thiesbrummel *et al.* J - V scans at four ageing times (0, 90, 280 and 480 min). Eight parameters were inferred: $\{N_0, D_I, d_E, d_H, v_n^E, v_p^E, v_n^H, v_p^H\}$, with all remaining model parameters fixed at educated estimates; because the inferred set is restricted, weakly identifiable parameters may shift to compensate for model mismatch, so trends in non-ionic parameters should be interpreted more cautiously than those in N_0 and D_I [22].

The most consistent ageing signature is ionic: the inferred mobile vacancy density N_0 increases by approximately three orders of magnitude over the ageing window (from $\sim 7 \times 10^{21} \text{ m}^{-3}$ to $\sim 2 \times 10^{25} \text{ m}^{-3}$), while the inferred ionic diffusion coefficient D_I decreases comparably (from $\sim 6 \times 10^{-13} \text{ m}^2 \text{ s}^{-1}$ to $\sim 5 \times 10^{-16} \text{ m}^2 \text{ s}^{-1}$). Physically, N_0 sets the amount of mobile ionic charge available to redistribute under bias, while D_I controls how quickly redistribution occurs relative to the voltage sweep [21, 22].

Under applied bias, mobile ion vacancies drift and accumulate near the transport-layer interfaces, forming space-charge (Debye) layers that partially screen the internal electric field in the perovskite. As N_0 increases this screening becomes stronger, flattening bands in the absorber and reducing field-assisted charge separation; this primarily impacts slow-scan or steady-state operation, while fast-scan (ion-frozen) performance is less affected. These trends therefore support the interpretation of Thiesbrummel *et al.* that ageing is dominated by increasingly strong ion-induced field screening rather than a primary degradation of electronic material quality [20].

The simultaneous decrease in D_I implies slower ionic motion and longer relaxation times, increasing sensitivity to scan rate and measurement history because the ionic distribution may not fully equilibrate during a given J - V sweep. Together, increasing N_0 and decreasing D_I describe an aged device with more mobile ions that respond more slowly, enhancing both internal field screening and non-equilibrium effects in the measured J - V curves.

Panels (e) and (f) show the joint posterior distributions for N_0 and D_I and reveal a strong, increasing correlation with ageing, which is expected because scan-rate-dependent J - V features

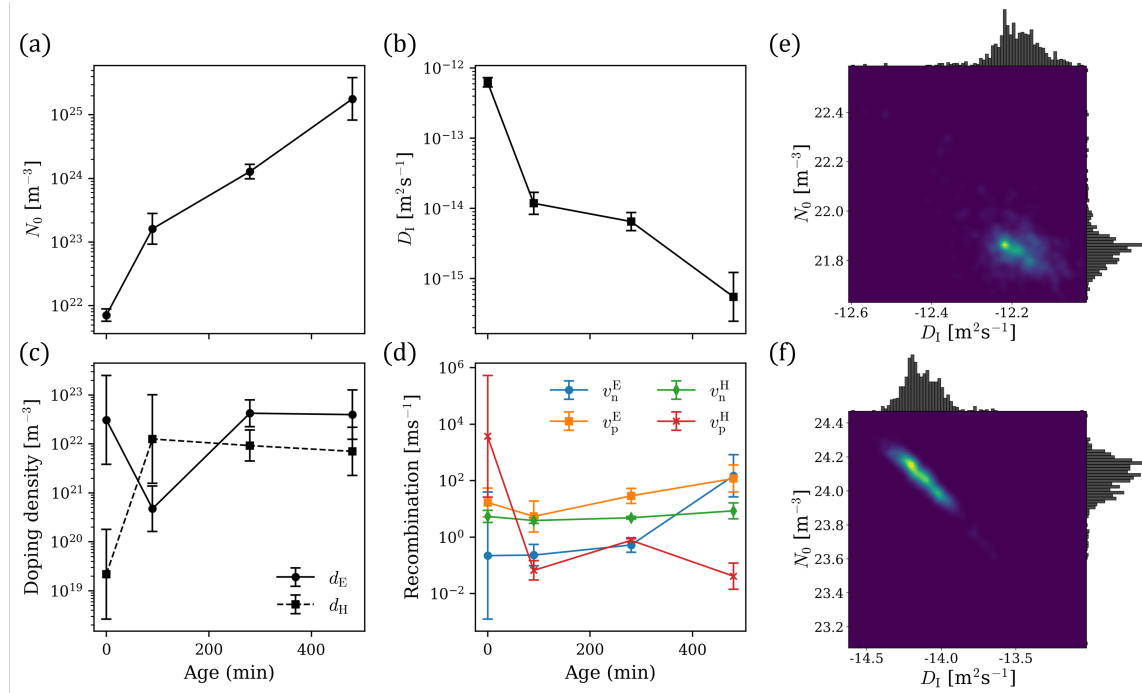


Figure 3. Bayesian parameter estimation results for the illumination-ageing series of Thiesbrummel *et al.* [20]. Panels (a)–(d) show the evolution of inferred parameters on a log scale with ageing for a pin-type triple-cation perovskite solar cell ($\text{Cs}_{0.05}(\text{FA}_{0.83}\text{MA}_{0.17})_{0.95}\text{Pb}(\text{I}_{0.83}\text{Br}_{0.17})_3$) aged under 1 sun illumination at open-circuit voltage. Error bars indicate 1σ credible intervals. Panels (e)–(f) show the joint posterior distributions for the mobile vacancy density N_0 and ionic diffusion coefficient D_I at 0 and 280 min.

constrain an effective ionic response that depends on both ionic charge magnitude (set by N_0) and timescale (set by D_I). Different combinations can therefore reproduce similar hysteresis signatures, concentrating posterior probability along a correlated trade-off direction [21, 22].

By contrast, the inferred transport-layer dopings and interfacial recombination velocities do not show a clear or monotonic trend across all ageing points; these variations are more consistent with overlap in parameter effects and disentanglement issues than with systematic changes in doping or recombination. Overall, the BPE results support the conclusion that the dominant effect of illumination ageing in this dataset is increased ion-induced internal field screening rather than severe electronic degradation at the interfaces or in the bulk absorber [20].

3.2. Comparison of squared errors: MCMC, SMC and ENN

Figure 4 compares parameter-recovery accuracy for Metropolis–Hastings MCMC, tempered SMC, and the ENN on a synthetic benchmark generated with IONMONGER. For each test case, IONMONGER is used to produce the multi-scan-rate J – V summary features from known parameters, and each inverse method is then applied to infer the selected parameters. Accuracy is quantified per-parameter using the squared relative error, $(1 - \hat{p}/p)^2$, shown on a logarithmic scale as box-and-whisker plots. Results are aggregated over 100 test samples for ENN and MCMC, and 5 for SMC (limited by runtime).

The dominant trend in Fig. 4 is that SMC provides the lowest errors for the most influential parameters in this setting: the mobile-ion quantities N_0 and D_I , and the transport-layer doping densities d_E and d_H . For these four parameters, the SMC error distributions are shifted to smaller values, indicating more reliable recovery of the true values under the same forward-model assumptions.

For the remaining parameters (the interfacial recombination velocities), the three methods perform similarly: their error distributions largely overlap, suggesting that these quantities are comparatively harder to identify from the J – V summary features used here and that, under this synthetic matched-model setting, no method offers a consistent advantage.

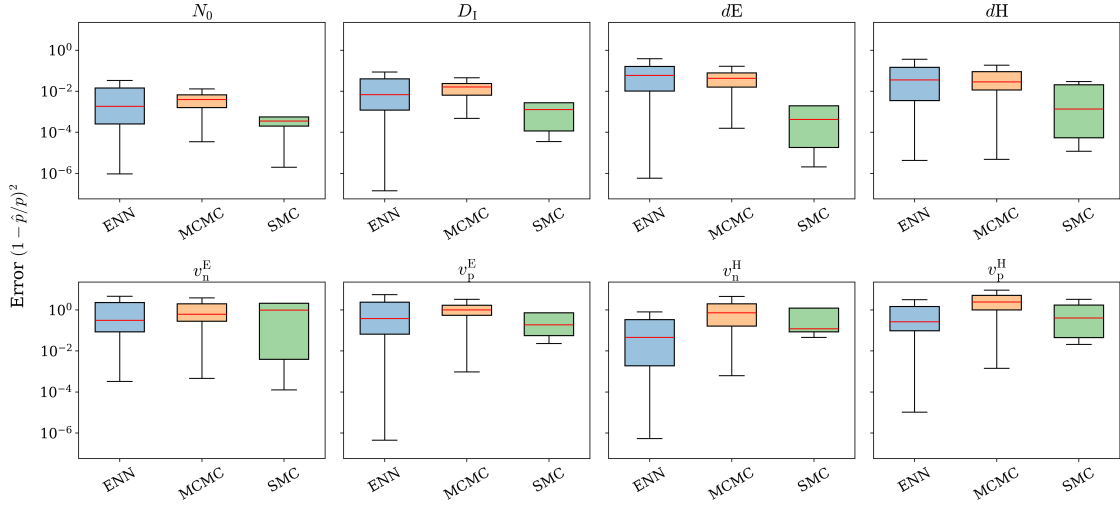


Figure 4. Box-and-whisker plots of the squared relative error, $(1 - \hat{p}/p)^2$, shown on a logarithmic scale for each model parameter. For each parameter, results from different inference methods are compared within the same subplot. Boxes denote the interquartile range with the median indicated by the central line, and whiskers represent the Tukey range ($1.5 \times \text{IQR}$). Individual samples are not shown. Errors are plotted on a logarithmic scale to highlight differences spanning multiple orders of magnitude. 100 samples were taken from ENN and MCMC and 5 from SMC.

While the ENN does not systematically outperform SMC on N_0 , D_I , dE , or dH , it does occasionally achieve markedly smaller errors than both sampling-based approaches for some parameters, visible as substantially lower-error cases in the distributions. This behaviour is consistent with the ENN learning highly informative mappings in certain regions of parameter space, even though its typical performance remains comparable to the samplers for parameters that are weakly constrained by the data. Overall, on synthetic IONMONGER data the SMC sampler is most accurate for the ionic and doping parameters, whereas all three approaches are broadly comparable for the remaining parameters, with the ENN sometimes providing notably more accurate point estimates on individual cases.

3.3. Comparison of measured and predicted impedance spectra for a solar cell

Figure 5 compares the complex impedance of a representative “cell A” spectrum (black, IONMONGER target) with spectra regenerated using parameters inferred by each inverse method. In a Nyquist plot, processes separate by frequency: at high frequencies the voltage oscillates too quickly for ions to move, so the response is dominated by electronic transport and displacement current, giving a compact arc commonly associated with recombination resistance in parallel with a geometric capacitance [27]; at lower frequencies mobile ion vacancies redistribute and accumulate at interfaces, modifying Debye layers and the internal potential and producing an additional low-frequency arc [21].

For this impedance experiment, seven parameters were inferred: $\{N_0, D_I, d_H, \varepsilon_E, \varepsilon_H, D_n, D_p\}$. The ionic parameters mainly control the low-frequency response: increasing N_0 increases the mobile charge available to screen the internal field and shifts the operating point about which the small-signal response is measured, while D_I sets the ionic relaxation time and therefore the frequency at which the ionic arc appears [21, 22]. The remaining parameters primarily affect the fast electronic response, with transport-layer dopings influencing band bending, carrier extraction and interfacial recombination, and permittivities contributing to the geometric capacitance and shifting the high-frequency reactance [27]. The diffusion coefficients were found to have very little effect on the graph output.

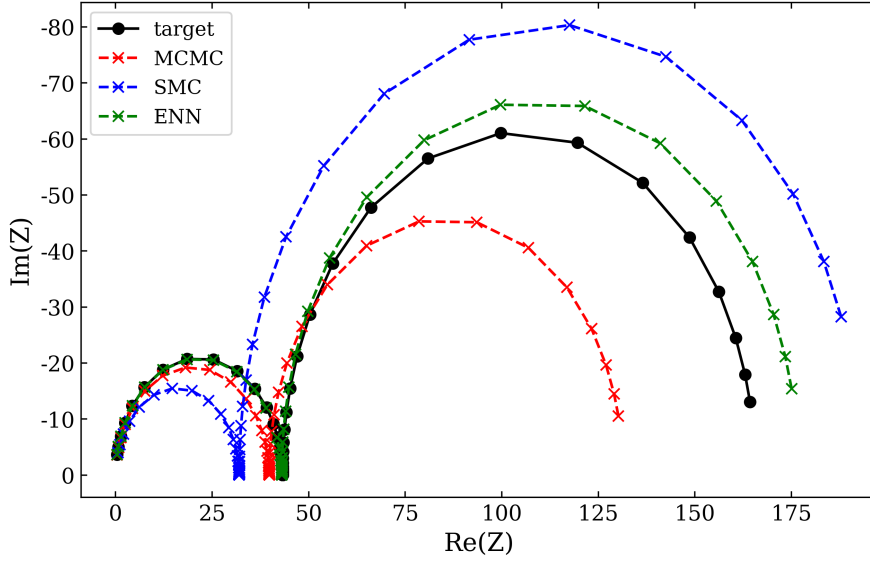


Figure 5. Comparison of impedance spectroscopy Nyquist plots for cell A produced using MCMC, SMC and ENN, compared to the IONMONGER target spectrum. Target data was generated using constants for cell A in Clarke et al. [27]

Consistent with these sensitivities, the largest differences between inferred spectra occur at low frequency: the ENN gives the closest overall match to the IONMONGER spectrum, SMC and MCMC shows the largest deviations. This is expected because the target spectrum was generated by IONMONGER using the same fixed constants employed during ENN training, so model mismatch is minimal, while sampling-based methods must explore a high-dimensional error surface (128 real and imaginary values across 64 frequencies) and are sensitive to parameter trade-offs because Nyquist arcs typically constrain combinations of resistances and capacitances rather than individual parameters [27].

The computational cost reflects the long simulation times required for low-frequency points: impedance inversions using SMC and MCMC take close to a day (SMC using $N = 500$ particles), whereas the ENN again produces parameter estimates in a few seconds once trained [27]. Possibly the SMC model could be improved by providing an alternative to MSE that deals with the graph as a whole rather than a point by point analysis.

3.4. ENN tracking of parameters

Figure 6 shows how the ENN responds when individual device parameters are varied across, and beyond, the ranges covered by the training data. Within the training range the ENN tracks the true parameter smoothly with low predicted uncertainty, indicating that the simulated $J-V$ and impedance features contain sufficient information to identify that parameter, while outside the training range the predicted uncertainty increases sharply, correctly signalling extrapolation.

This behaviour is most informative for parameters that strongly affect the electrical response, particularly N_0 and D_I , which control the amount and timescale of ionic charge accumulation at interfaces and thereby modify charge extraction through field screening [20, 21]. In impedance spectroscopy, these same ionic parameters determine the appearance and position of the low-frequency arc, while geometric and electronic parameters mainly affect the high-frequency response [27]; the ENN can therefore serve as a rapid screening tool, where low uncertainty indicates operation within the trained regime and high uncertainty flags unusual behaviour or missing physics.

Figure 6 also highlights a limitation of the current ENN approach: unlike sampling-based BPE methods, which can be rerun with different priors or updated model assumptions, the ENN is constrained by the parameter ranges and fixed constants used in training. This can reduce accuracy

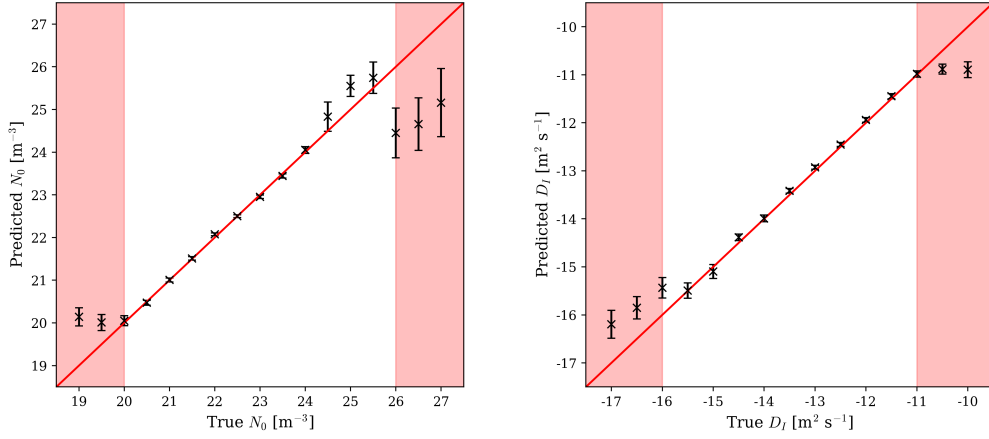


Figure 6. ENN predictions of device parameters and associated uncertainties are shown for inputs spanning the training distribution and extending into out-of-distribution regimes. The red shaded regions denote parameter values outside the training ranges specified in Table 2. Within the trained domain, the ENN produces estimates with high accuracy and precision. Outside this range, prediction accuracy degrades and the estimated epistemic uncertainty increases, while aleatoric uncertainty remains negligible. The parameters exhibiting the lowest squared error in Fig. 4 are shown on a log scale: (a) N_0 and (b) D_1 .

under model mismatch; improving robustness for experimental devices likely requires training datasets that span a wider range of device parameters and include uncertainty in typically fixed quantities such as layer thicknesses, illumination intensity and contact properties [20, 22].

Overall, these results show that the most suitable inverse method depends on both the data type and forward-model fidelity: for experimental J - V data, where uncertainty in device constants is unavoidable, SMC provides the most accurate and stable parameter estimates, while for synthetic impedance spectra generated under the same assumptions as training, the ENN offers the highest accuracy at a fraction of the computational cost. A practical strategy is to use the ENN for rapid initial analysis and uncertainty screening, followed by targeted SMC inference when higher accuracy or more detailed physical interpretation is required [20, 22].

4. Conclusion

Using IONMONGER as a physics-based forward model, this report compared tempered Sequential Monte Carlo (SMC) Bayesian inference with an evidential neural network (ENN) surrogate for extracting mobile-ion parameters from J - V and impedance data. For the illumination-ageing J - V series, SMC gave the most reliable fits and showed a clear ionic trend: the mobile vacancy density N_0 increased by $\sim 10^3$ while the ionic diffusion coefficient D_1 decreased by $\sim 10^3$. This implies an aged device has more ionic charge that redistributes more slowly, leading to stronger interfacial Debye layers that screen the internal field; this mainly harms slow-scan and steady-state performance, supporting the discussion conclusion that the observed PCE loss is driven primarily by ion-induced field screening rather than a clear, systematic worsening of electronic/interface parameters. The joint posteriors also show a strong N_0 - D_1 trade-off, explaining why an ensemble method like SMC is robust on this dataset. For impedance spectroscopy under matched (synthetic) assumptions, the ENN reproduced the target spectrum most closely and reduced inference time from nearly a day to seconds, but its accuracy drops when experimental devices differ from the fixed constants and ranges used in training. With further development – in particular training the ENN across a wider range of device conditions and parameters (e.g., thickness, illumination, and contacts) and replacing pointwise MSE with a spectrum-aware impedance objective – the ENN could progress from a fast screening tool to providing near-instant analysis of PSCs and their degradation pathways, with SMC reserved for high-confidence interpretation under model mismatch.

5. Acknowledgments

I would like to thank our supervisors, Prof. Alison Walker and Dr. Kjeld Jensen, for their assistance and support, providing deep knowledge of physics/statistics. I would also like to thank my project partner Ramin Uddin who produced the ENN.

References

- [1] Hayat M B, Ali D, Monyake K C, Alagha L Z and Ahmed N 2019 *Int. J. Energy Res.* **43** 1049–1067
- [2] Guney M S 2016 *Renew. Sustain. Energy Rev.* **57** 776–785
- [3] Green M A, Levi D H, Hishikawa Y, Dunlop E D, Hohl-Ebinger J and Ho-Baillie A 2018 *Prog. Photovolt: Res. Appl.* **26** 3–12
- [4] Lee S W, Bae S, Kim D and Lee H S 2020 *Adv. Mater.* **32** 2002202
- [5] Nayak P K, Mahesh S, Snaith H J and Cahen D 2019 *Nat. Rev. Mater.* **4** 269–285
- [6] Kojima A, Teshima K, Shirai Y and Miyasaka T 2009 *J. Am. Chem. Soc.* **131** 6050–6051
- [7] Wehrenfennig C, Eperon G E, Johnston M B, Snaith H J and Herz L M 2014 *Adv. Mater.* **26** 1584–1589
- [8] Snaith H J 2013 *J. Phys. Chem. Lett.* **4** 3623–3630
- [9] National Renewable Energy Laboratory 2024 Best research-cell efficiency chart available at <https://www.nrel.gov/pv/cell-efficiency.html>
- [10] Green M A, Ho-Baillie A and Snaith H J 2014 *Nat. Photonics* **8** 506–514
- [11] Yang C *et al.* 2024 *Light Sci. Appl.* **13** 227
- [12] Eperon G E, Hörantner M T and Snaith H J 2017 *Nat. Rev. Chem.* **1** 0095
- [13] Park N G 2015 *Mater. Today* **18** 65–72
- [14] Boyd C C, Cheacharoen R, Leijtens T and McGehee M D 2019 *Chem. Rev.* **119** 3418–3451
- [15] Duan L *et al.* 2023 *Nat. Rev. Mater.* **8** 261–281
- [16] Richardson G *et al.* 2016 *Energy Environ. Sci.* **9** 1476–1485
- [17] Calado P, Telford A M, Bryant D, Li X, Nelson J, O'Regan B C and Barnes P R F 2016 *Nat. Commun.* **7** 13831
- [18] Guerrero A *et al.* 2014 *Appl. Phys. Lett.* **105** 133902
- [19] Futscher M H, Lee J M, McGovern L, Muscarella L, Amassian A, van Ruitenbeek E J W and Abate A 2019 *Mater. Horiz.* **6** 1497–1503
- [20] Thiesbrummel J *et al.* 2024 *Nat. Energy* **9** 664–676
- [21] Cave J M *et al.* 2020 *J. Appl. Phys.* **128** 184501
- [22] McCallum S G, Nicholls O, Jensen K O, Cowley M V, Lerpinière J E and Walker A B 2024 *J. Phys. Energy* **6** 015005
- [23] Guerrero A, Juarez-Perez E J, Todorov N K, Edri S, Koirala S P, de Wolf S, Budden W P, Sessolo M, Segawa H, Abate A and Bisquert J 2016 *J. Phys. Chem. C* **120** 8023–8032
- [24] Tennyson E M, Neal B A and Cotsworth J M 2019 *J. Phys. Energy* **1** 012001
- [25] Courtier N E, Walker A B and Richardson G 2019 *J. Comput. Electron.* **18** 1081–1090
- [26] Clarke W, Bennett L J, Grudeva Y, Foster J M, Richardson G and Courtier N E 2023 *J. Comput. Electron.* **22** 364–382
- [27] Clarke W, Cameron P and Richardson G 2024 *J. Phys. Chem. Lett.* **15** 11730–11736
- [28] Roberts G O and Rosenthal J S 2001 *Statist. Sci.* **16** 351–367
- [29] Moral P D, Doucet A and Jasra A 2006 *J. R. Stat. Soc. B* **68** 411–436
- [30] Chopin N and Papaspiliopoulos O 2020 Smc samplers *An Introduction to Sequential Monte Carlo* Springer Series in Statistics (Springer) pp 329–353
- [31] Zhou Y, Johansen A M and Aston J A D 2016 *J. Comput. Graph. Stat.* **25** 701–726
- [32] Wang S, Ge S, Doig R and Wang L 2022 *J. Comput. Graph. Stat.* **31** 600–613
- [33] Garthwaite P H, Fan Y and Sisson S A 2016 *Commun. Stat. Theory Methods* **45** 5098–5111
- [34] Spencer S E F 2021 *Aust. N. Z. J. Stat.* **63** 468–484
- [35] Robbins H and Monro S 1951 *Ann. Math. Stat.* **22** 400–407
- [36] Gao J, Chen M, Xiang L and Xu C 2024 *arXiv preprint arXiv:2409.04720*
- [37] Jøsang A 2001 *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **9** 279–311
- [38] Jøsang A 2016 *Subjective Logic* (Springer)
- [39] Amini A, Schwarting W, Soleimany A and Rus D 2020 Deep evidential regression *Advances in Neural Information Processing Systems* vol 33 pp 14927–14937
- [40] Kiureghian A D and Ditlevsen O 2009 *Struct. Saf.* **31** 105–112