Population Synthesis as Scenario Generation for Simulation-based Planning under Uncertainty

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

Arnau Quera-Bofarull
University of Oxford

Nicholas Bishop
University of Oxford

J. Doyne Farmer
University of Oxford

Anisoara Calinescu
University of Oxford

Michael Wooldridge
University of Oxford

ABSTRACT

Agent-based models have the potential to become instrumental tools in real-world decision-making, equipping policy-makers with the ability to experiment with high-fidelity representations of complex systems. Such models often rely crucially on the generation of synthetic populations with which the model is simulated, and their behaviour can depend strongly on the population's composition. Existing approaches to synthesising populations attempt to model distributions over agent-level attributes on the basis of data collected from a real-world population. Unfortunately, these approaches are of limited utility when data is incomplete or altogether absent – such as during novel, unprecedented circumstances – so that considerable uncertainty regarding the characteristics of the population being modelled remains, even after accounting for any such data. What is therefore needed in these cases are tools to simulate and plan for the possible future behaviours of the complex system that can be generated by populations that are consistent with this remaining uncertainty. To this end, we frame the problem of synthesising populations in agent-based models as a problem of *scenario generation*. The framework that we present is designed to generate synthetic populations that are on the one hand consistent with any persisting uncertainty, while on the other hand matching closely a target, user-specified scenario that the decision-maker would like to explore and plan for. We propose and compare two generic approaches to generating synthetic populations that produce target scenarios, and demonstrate through simulation studies that these approaches are able to automatically generate synthetic populations whose behaviours match the target scenario, thereby facilitating simulation-based planning under uncertainty.

KEYWORDS

agent-based models; scenario planning; synthetic data

ACM Reference Format:
Joel Dyer, Arnau Quera-Bofarull, Nicholas Bishop, J. Doyne Farmer, Anisoara Calinescu, and Michael Wooldridge. 2024. Population Synthesis as Scenario Generation for Simulation-based Planning under Uncertainty. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 9 pages.

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

1 INTRODUCTION

Agent-based models (ABMs) are computational models consisting of a population of agents – which may, for example, represent individuals or households in a population, and/or firms in an economy – interacting with each other and their environment to simulate the underlying complex system in a bottom-up fashion. Such models are a promising tool for modelling complex systems across a variety of domains, from economics [2, 8] to epidemiology [36].

By defining the individual behaviour of each agent, one hopes that macroscopic properties of interest will emerge naturally as a consequence of the collective actions of agents. By modelling the system of interest at this level of granularity, the modeller can draw insights into the ways in which individual behaviours and actions give rise to emergent phenomena at the macroscopic level. For example, in epidemiology, agent-based models can help to explain how the spread of disease is related to the actions of individuals, allowing for the design of effective policies at the individual level which limit infection. For this reason, ABMs are lauded as a promising tool for performing "what-if" scenario analyses and experimenting with policy interventions in the simulated environments they provide, permitting modellers and decision-makers to go beyond purely predictive analyses and instead plan for various possible future scenarios and potential outcomes by interacting with the simulation model.

Population synthesis – the process of generating the synthetic populations with which a given ABM may be simulated forwards – is a key methodological component of the science of agent-based modelling, given the central role synthetic populations play in the specification of these models. In the ABM literature, the problem of population synthesis has in many cases been interpreted as the problem of finding a distribution $\iota(\mathbf{z} \mid \boldsymbol{\theta})$, parameterised by parameters $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^d$, over agent attributes $\mathbf{z} \in \mathcal{Z} \subseteq \mathbb{R}^a$ that – according to some measure of congruence – matches the distribution $g(\mathbf{z})$ of attributes measured in the real-world counterpart to a given synthetic population [9, 10, 34]. Here, Θ and $\mathcal{Z}$ denote the d- and a-dimensional spaces of values that the parameters $\boldsymbol{\theta}$ and agent-level attributes $\mathbf{z}$ can assume. In such cases, the parameters $\boldsymbol{\theta}$ are typically estimated using a finite sample of attributes seen in the real-world population under consideration, for example as collected by censuses [45]. Once this distribution over attributes is constructed, synthetic populations may be generated and used in downstream modelling tasks by sampling the attributes of individual agents independently from $\iota(\cdot \mid \boldsymbol{\theta})$ and using these samples to initialise each agent in the simulated population.

1.1 Limitations of existing approaches to population synthesis

Existing approaches to population synthesis:

(1) assume that there exists a sample of individual-level attributes from a real-world population to be modelled; and
(2) perform population synthesis upfront, such that the population synthesis procedure is not informed by the behaviour of the ABM under the generated population.

Such approaches to population synthesis suffer from important limitations as a consequence of these two properties. Given that the appeal of agent-based simulation derives partly from the opportunity it provides modellers and decision-makers to experiment with, explore, and plan for possible future scenarios in complex systems, there is a need for population synthesis approaches that entail the generation of populations that are not yet necessarily realised in the current state of the real world, and/or for which data does not exist or is not accessible. This would better equip modellers and decision-makers with the prerequisite tools for planning and experimentation using ABMs, rather than limiting the experimenter to the state of the world as it currently stands. In these cases, there may be no samples from any real-world population through which a distribution over agent-level attributes can be learned and from which a suitable synthetic population may be generated. This may also be the case even in situations where it is desirable for the synthetic population to mimic a real-world population on which data may possibly be gathered in principle, but for which data-gathering may be difficult or impossible in practice due to e.g., expense, time constraints, or privacy-related concerns. Existing approaches to population synthesis for ABMs are incompatible with such situations, given the first feature of these approaches listed above. When this is true, alternative approaches to population synthesis that rely not on any data sampled from any real-world population, but rather on the scenario to be planned for, would be advantageous. However, existing approaches to population synthesis for ABMs are once again not designed to generate specific scenarios, but are instead designed to model the distribution of attributes in a sample from the population.

1.2 A concrete motivating example

The COVID-19 pandemic caused severe disruption globally. As the emergency began to unfold in late 2019 and early 2020, there was considerable uncertainty regarding the range of factors driving the pandemic and subsequent economic fallout, including: the degree to which individuals would comply with state-imposed social and travel restrictions; the readiness with which consumers would panic-buy and hoard goods in limited supply; and the influence of vaccine-scepticism on treatment uptake. Policy- and decision-makers in all aspects of life were faced with the task of planning for the future of their businesses, households, and countries in the unprecedented circumstances with which they were presented and under these conditions of considerable uncertainty. In such cases, constructing and exploring the behaviour of simulations of the developing situation – such as through ABMing – can enable these decision-makers to understand the possible impacts of this uncertainty, develop intuition, and reason about the possible future trajectories that could materialise.

To make effective use of such simulations, the decision-maker might desire the capability to generate specific scenarios within the simulated environment, which would allow them to prepare effective, scenario-specific response strategies. In ABMs and other complex systems, the readiness with which different scenarios can be realised or simulated may depend strongly on the composition of the synthetic population used within the simulation. Given the novelty of the circumstances brought forward by the COVID-19 pandemic, key parameters and characteristics of the population as they relate to the progression of the pandemic were unknown (such as the susceptibility of its constituent citizens to panic-buying, the readiness with which they will conform with social restrictions, and the prevalence of vaccine-scepticism mentioned above). Thus, it is crucial to develop the ability to rapidly experiment with different settings for these population-level parameters and the scenarios they generate in order to facilitate effective decision-making in this and other situations characterised by uncertainty. This ability will, however, rely on the ability to identify population-level parameters that generate different scenarios, which motivates our contribution.

1.3 Our contribution

Based on our discussion in Subsection 1.1, existing approaches to population synthesis for ABMs are not designed for situations such as in the motivating example presented in Subsection 1.2. To address those shortcomings of existing approaches to population synthesis, we consider the problem of synthesising populations that give rise to a desired target scenario, answering the question:

> "*Under the assumption that the model is accurate, what might the population in this system need to look like in order to realise a user-specified scenario?*"

We formulate this problem mathematically and discuss several potential approaches and solutions. Our resultant framework for population synthesis for ABMs relies on the ability of the user to specify a loss function determining the proximity of the model's behaviour to a desired target scenario, in addition to a parametric family of distributions from which agent attributes may be generated. Finally, we provide an open-source software package[1] implementing the solutions we discuss, with which we perform and present simulation studies to illustrate our proposed framework for joint population synthesis and scenario generation in ABMs. We believe that our contribution will facilitate the use of ABMs as an environment for experimentation and performing scenario analyses, and will enable decision-makers to explore, reason about, and plan for the possible behaviours of a complex system under uncertainty regarding the properties of the system's population.

2 PROBLEM STATEMENT

In this section, we formulate the problem of population synthesis for ABMs mathematically and establish essential notation. Let $\mathbf{x} \in \mathcal{X}$ denote the full output time series, specifying the states of all agents in the ABM over all time steps $t = 1, \ldots, T$, where $\mathcal{X}$ is the space of possible values $\mathbf{x}$ can assume. We assume that the ABM consumes structural parameters $\boldsymbol{\omega} \in \Omega \subseteq \mathbb{R}^w$, $w \geq 1$ and a collection of agent-level attributes $\mathcal{A}_N := (\mathbf{z}_n)_{n=1,\ldots,N}$, where $N \geq 1$ is the number

[1] Available at https://github.com/joelnmdyer/synthpop.

Algorithm 1: Scenario generation in an ABM through population synthesis.

Data: Proposal distribution q
Result: Simulations $\{\mathbf{x}^{(r)}\}_{r=1}^{R}$ matching the target scenario
for $r \in \{1, \ldots, R\}$ **do**
  Sample $(\boldsymbol{\omega}^{(r)}, \boldsymbol{\theta}^{(r)}) \sim q(\cdot)$;
  Sample attributes for N agents: $\mathcal{A}_n^{(r)} \sim f(\cdot \mid \boldsymbol{\theta}^{(r)})$;
  Forward simulate from the agent-based simulator:
   $\mathbf{x}^{(r)} \sim p(\cdot \mid \boldsymbol{\omega}^{(r)}, \mathcal{A}_N^{(r)})$
end

of agents in the system, before performing a stochastic forward run such that $\mathbf{x} \sim p(\cdot \mid \boldsymbol{\omega}, \mathcal{A}_N)$, where $p(\cdot \mid \boldsymbol{\omega}, \mathcal{A}_N)$ is the ABM's likelihood function. In addition, we assume that the collection $\mathcal{A}_N$ is jointly generated from a distribution $f(\cdot \mid \boldsymbol{\theta})$ with parameters $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^d$. Here, $\boldsymbol{\theta}$ may be seen as population-level parameters, which probabilistically determine the attributes of the individuals comprising the synthetic population.

Our goal is to devise approaches for automatically and reliably identifying values for $\boldsymbol{\theta}$ that synthesise populations of agents that tend to realise the simulated scenario of interest to the experimenter. In other words, the goal is to generate parameters for an ABM that generates a population with a desired behaviour or behaviours. In general, we might like to simultaneously vary and identify the corresponding values of $\boldsymbol{\omega}$. The problem we consider is therefore the problem of generating a distribution q over the joint space $\mathcal{P} := \Omega \times \Theta$, in which high probability mass is assigned to regions of $\mathcal{P}$ that provide the best matches to the desired scenario, while the regions of $\mathcal{P}$ that produce a comparatively poor match to the desired scenario are assigned a low probability mass. Having identified such a distribution q for this hierarchical model, a diverse range of simulations that best match the desired scenario may be generated from the agent-based simulator by sampling from the distribution

$$p(\mathbf{x}) = \int_{\Omega \times \Theta} p(\mathbf{x} \mid \boldsymbol{\omega}, \boldsymbol{\theta})\, q(\boldsymbol{\omega}, \boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\omega}\, \mathrm{d}\boldsymbol{\theta}, \tag{1}$$

where

$$p(\mathbf{x} \mid \boldsymbol{\omega}, \boldsymbol{\theta}) = \int_{\mathcal{Z}^N} p(\mathbf{x} \mid \boldsymbol{\omega}, \mathcal{A}_N)\, f(\mathcal{A}_N \mid \boldsymbol{\theta})\, \mathrm{d}\mathcal{A}_N.$$

This sampling procedure is presented schematically in Figure 1 and operationally in Algorithm 1. In the special case that the agent attribute vectors $\mathbf{z}_n$ are all conditionally independent given $\boldsymbol{\theta}$, we have the following factorisation:

$$f(\mathcal{A}_N \mid \boldsymbol{\theta}) = \prod_{n=1}^{N} \iota(\mathbf{z}_n \mid \boldsymbol{\theta}),$$

where the $\mathbf{z}_n \overset{\text{iid}}{\sim} \iota(\cdot \mid \boldsymbol{\theta})$, $n = 1, \ldots, N$ for some distribution $\iota(\cdot \mid \boldsymbol{\theta})$.

3 METHODS

Given the problem specification constructed in Section 2, the technical challenge lies in obtaining a suitable proposal distribution q. We discuss two such possibilities here. These approaches require that the experimenter specifies a loss function $\ell : \mathcal{X} \to [0, \infty)$

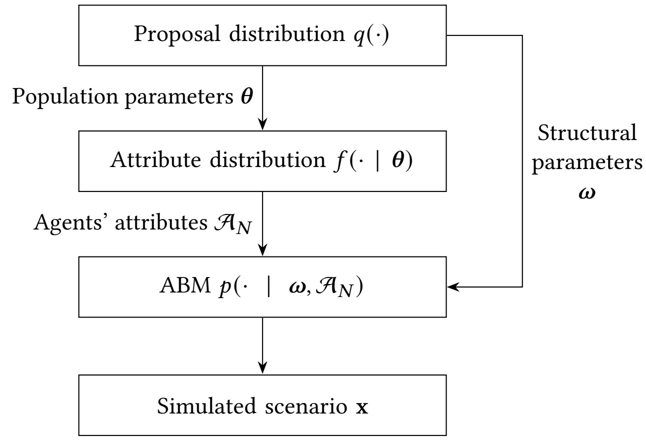

Figure 1: Schematic illustration of Algorithm 1.

(which we assume to be non-negative) that quantifies the similarity between a simulated output $\mathbf{x}$ from the ABM and the target scenario the experimenter would like to generate. Furthermore, in what follows we use the notation

$$\mathcal{L}(\boldsymbol{\omega}, \boldsymbol{\theta}) = \mathbb{E}_{p(\mathbf{x} \mid \boldsymbol{\omega}, \boldsymbol{\theta})} \left[h_\epsilon\left(\ell(\mathbf{x})\right) \right],$$

where h_ϵ is a method-dependent function parameterised by ϵ.

3.1 Threshold-based sampling (TBS)

In this first class of methods we consider, we propose to down-weight candidate values for the parameters $(\boldsymbol{\omega}, \boldsymbol{\theta})$ by letting h_ϵ be a probability kernel function with hyperparameter $\epsilon > 0$, before letting

$$q(\boldsymbol{\omega}, \boldsymbol{\theta}) \propto \mathcal{L}(\boldsymbol{\omega}, \boldsymbol{\theta}). \tag{2}$$

For example, we may choose

$$h_\epsilon(\cdot) \propto \mathbb{I}(\cdot \leq \epsilon), \tag{3}$$

which evaluates to 1 if $\ell \leq \epsilon$ and 0 if $\ell > \epsilon$. This corresponds to

$$q(\boldsymbol{\omega}, \boldsymbol{\theta}) \propto \mathbb{P}\left(\{\ell(\mathbf{x}) \leq \epsilon \mid \mathbf{x} \sim p(\cdot \mid \boldsymbol{\omega}, \boldsymbol{\theta})\}\right),$$

meaning that any given combination of structural and population-level parameter, $(\boldsymbol{\omega}, \boldsymbol{\theta})$, is down-weighted if the probability with which it results in populations that produce a scenario that is within an ϵ-ball of the desired scenario is comparatively low. Other related alternatives are available however; a simple example is the Gaussian kernel, $h_\epsilon(\cdot) \propto \exp(- \cdot / \epsilon)$, which performs a similar but softer action to the indicator function used previously. In either case, a set $\mathcal{T}_I := \{\boldsymbol{\omega}^{(i)}, \boldsymbol{\theta}^{(i)} \mid i = 1, \ldots, I\}$ of $I \geq 1$ samples can be generated from the resultant proposal distribution q in a Monte Carlo fashion, for example with rejection sampling [44], Markov chain Monte Carlo [22, 32], or sequential Monte Carlo [4, 30]. With the samples $\mathcal{T}_I$, a range of forward simulations from the ABM that closely match the scenario to be generated can be produced by synthesising populations at the pairs $(\boldsymbol{\omega}^{(i)}, \boldsymbol{\theta}^{(i)})$, $i = 1, \ldots, I$ and forward simulating the model, such that for all i,

$$\mathbf{x}^{(i)} \sim p(\cdot \mid \boldsymbol{\omega}^{(i)}, \mathcal{A}_N^{(i)}), \quad \mathcal{A}_N^{(i)} \sim f(\cdot \mid \boldsymbol{\theta}^{(i)}).$$

We note that, in general, the value of the hyperparameter ϵ will determine the variance of the proposal distribution q, with smaller values resulting in a proposal distribution that is more concentrated in regions of $\mathcal{P}$ that synthesise populations with the specified scenario-generating tendencies. This, however, will tend to reduce the variety in the different scenarios generated by the ABM, amounting to a trade-off between (a) an exploration of the different scenarios that can be generated by a synthetic population that bear a sufficiently close resemblance to the target scenario on the one hand, and (b) an exploitation of the regions of $\mathcal{P}$ that generate the best matches to the target scenario, as specified by ℓ, on the other hand. We note that the experimenter may choose to pick larger values of ϵ when unsure about the correct specification of the loss function to generate the target scenario in mind. Finally, we herein term this approach threshold-based sampling, abbreviated as TBS.

3.2 Variational optimisation (VO)

A second class of methods we consider involves constructing the proposal distribution q by solving a variational optimisation (VO) problem. In particular, we let h_ϵ be the identity function and $\boldsymbol{Q} := \{q(\cdot \mid \boldsymbol{\phi}) \mid \boldsymbol{\phi} \in \Phi\}$ be a family of probability distributions indexed by parameter $\boldsymbol{\phi}$ whose values lie in some set Φ. We then find q as the solution to the optimisation problem

$$q = \arg\min_{\phi \in \Phi} \left\{ \mathbb{E}_{\boldsymbol{\omega},\boldsymbol{\theta} \sim q(\boldsymbol{\omega},\boldsymbol{\theta}|\boldsymbol{\phi})} \left[\mathcal{L}(\boldsymbol{\omega},\boldsymbol{\theta})\right] - \gamma \cdot \mathbb{H}(q(\cdot \mid \boldsymbol{\phi})) \right\}, \quad (4)$$

where $\mathbb{H}(q) = \mathbb{E}_{\boldsymbol{\varrho} \sim q}\left[-\log q(\boldsymbol{\varrho})\right]$ is the (differential) entropy of q and $\gamma \geq 0$ is a regularisation hyperparameter. Setting $\gamma = 0$ will permit q to collapse into a degenerate distribution whose mass is concentrated around the pairs $(\boldsymbol{\omega}, \boldsymbol{\theta})$ that minimise $\mathcal{L}(\boldsymbol{\omega}, \boldsymbol{\theta})$. In contrast, larger values of γ will encourage greater variance in q, leading to greater diversity in the synthesised populations and, consequently, the scenarios that are generated by the simulator. In this way, γ plays a similar role to ϵ in TBS (see Section 3.1).

For VO, we assume that the $q(\cdot \mid \boldsymbol{\phi}) \in \boldsymbol{Q}$ are differentiable with respect to $\boldsymbol{\phi}$, such that they easily permit the construction of the distribution's score function $\nabla_{\boldsymbol{\phi}} \log q(\boldsymbol{\omega}, \boldsymbol{\theta} \mid \boldsymbol{\phi})$. This permits the use of gradient-based optimisation [7] to minimise the objective function in Equation (4) using a score-based Monte Carlo estimate

$$\frac{1}{R}\sum_{r=1}^{R} \left\{ \ell(\mathbf{x}^{(r)}) + \gamma \cdot l\left(\boldsymbol{\omega}^{(r)}, \boldsymbol{\theta}^{(r)} \mid \boldsymbol{\phi}\right) \right\} \cdot \nabla_{\boldsymbol{\phi}} l\left(\boldsymbol{\omega}^{(r)}, \boldsymbol{\theta}^{(r)} \mid \boldsymbol{\phi}\right)$$

of the gradient of the objective, where $l = \log q(\boldsymbol{\omega}, \boldsymbol{\theta} \mid \boldsymbol{\phi})$,

$$\left(\mathbf{x}^{(r)}, \boldsymbol{\omega}^{(r)}, \boldsymbol{\theta}^{(r)}\right) \sim p(\mathbf{x} \mid \boldsymbol{\omega}, \boldsymbol{\theta})\, q(\boldsymbol{\omega}, \boldsymbol{\theta} \mid \boldsymbol{\phi}), \quad r = 1, \ldots, R,$$

and where we have used the fact that $\mathbb{E}_{\boldsymbol{\varrho} \sim q(\cdot|\boldsymbol{\phi})}\left[\nabla_{\boldsymbol{\phi}} \log q(\boldsymbol{\varrho} \mid \boldsymbol{\phi})\right] = 0$ for any $q(\cdot \mid \boldsymbol{\phi})$. Alternatively, when sampling from the proposal $q(\cdot \mid \boldsymbol{\phi})$ can be reparameterised, lower-variance estimates of the gradient of the entropy regularisation term can alternatively be obtained with pathwise gradients [33].

4 SIMULATION STUDIES

In this section, we present simulation studies on two agent-based models using the approaches introduced in Section 3. While there are many possible implementations of TBS and VO, we consider the following specific example implementations in our experiments:

- For TBS we construct a proposal density q as in Equation (2) with h_ϵ taken to be a Uniform kernel as in Equation (3), and generate samples of $(\boldsymbol{\omega}, \boldsymbol{\theta})$ pairs from this proposal using a sequential Monte Carlo (SMC) sampling procedure. We refer to this instantiation of TBS as TBS-SMC.
- For VO, we take $q(\boldsymbol{\omega}, \boldsymbol{\theta} \mid \boldsymbol{\phi})$ to be an autoregressive normalising flow [38] with trainable parameters $\boldsymbol{\phi}$ optimised as described in Subsection 3.2. In the case that components of $(\boldsymbol{\omega}, \boldsymbol{\theta})$ are constrained to lie within open or closed intervals, samples from the flow are restricted to these domains with sigmoid transformations (see [16] for further details). We refer to this instantiation of VO as VO-NF.

Throughout, we fix a simulation budget of 10^5 simulations for TBS-SMC, and train VO-NF until the simulation budget is reached, unless convergence is achieved prior to this. Further details on TBS-SMC and VO-NF are given in the supplement [16].

4.1 Axtell's model of the emergence of firms

We consider a variation of the ABM of firm emergence in labour markets proposed in Axtell [1], in which the movement of agents between existing firms, or from existing firms to new firms, is simulated over time. In this model, each agent $n \in \{1, \ldots, N\}$ in the population of N agents is a labourer in an economy and works with some effort level $e_n^t \in [0, 1]$ at time $t \in [0, 1]$. At any given time t, agent n belongs to firm $\mathrm{f}_n^t \in \{1, \ldots, N\}$, where the model is initialised with $\mathrm{f}_n^0 = n$ (i.e., there are initially N firms, each being singleton sets). Broadly, each agent periodically reevaluates their situation in the labour market at some agent-specific characteristic rate $\rho_n \in \mathbb{R}_{\geq 0}$, which indexes a Poisson process for that agent. At each reevaluation event (each of which occurs in continuous time), the contemplative agent considers the trade-off – modelled with a Cobb-Douglas function and determined by the agent-specific parameter $\nu_n \in [0, 1]$ (see [16]) – between the utility they derive from (i) participating in firm f_n^t and sharing in its output, and (ii) the disutility experienced from contributing to the firm's productive activities at their current effort level e_n^t. High (resp. low) values of ν_n correspond to preference for income (resp. leisure). On this basis, the agent decides to either (i) readjust its effort levels, (ii) join a neighbouring firm, or (iii) start a new firm. Further details on the model are provided in the supplement [16].

The agents comprising the model have associated with them a set of three attributes – their initial effort levels e_n^0, the rate at which they reevaluate their positions ρ_n, and their relative preference for income and leisure ν_n – giving $\mathbf{z}_n = (e_n^0, \rho_n, \nu_n)$. Additionally, the utility an agent experiences by participating in any given firm is parameterised by structural parameters $\boldsymbol{\omega} = (a, b, \beta)$ (see [16]), which we also vary in the scenario generation process.

4.1.1 Can an initially hard-working population become lazy over time? The scenarios we consider and aim to generate with this model are scenarios in which initially hard-working populations (characterised by high effort levels) become lazy (characterised by low effort levels) over time. A suitable non-negative loss function ℓ for targeting this scenario is the shifted difference in average effort exerted by agents at the end and beginning of the simulated period,

$$\ell(\mathbf{x}) = \frac{1}{N}\sum_{n=1}^{N}\left(e_n^1 - e_n^0 + 1\right), \quad (5)$$

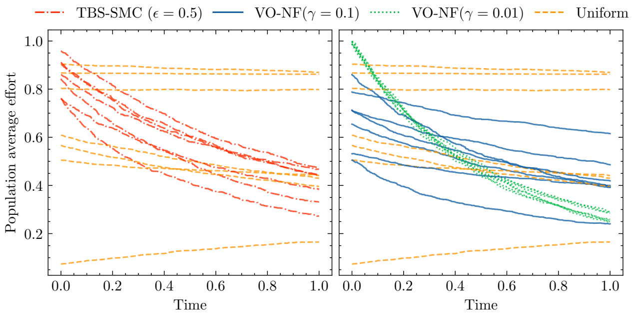

Figure 2: Simulations of population average effort over time generated from the marginal likelihood function (Equation (1)) obtained from (a) TBS-SMC (red), (b) VO-NF (blue and green), and (c) a uniform proposal over $\mathcal{P}$ (Uniform, orange).

in which $\mathbf{x}$ is the time series of all agents' effort levels at each reevaluation event, and the unit offset ensures the loss takes values in the range $[0, 2]$. To generate such scenarios, we take $f(\mathcal{A}_N \mid \boldsymbol{\theta}) = \prod_{n=1}^{N} \iota(\mathbf{z}_n \mid \boldsymbol{\theta})$ with

$$\iota(\mathbf{z}_n \mid \boldsymbol{\theta}) = \text{Beta}(e_n^0 \mid \varepsilon_a, \varepsilon_b) \cdot \text{Beta}(\nu_n \mid g_a, g_b) \cdot \text{Gamma}(\rho_n \mid \varrho_a, \varrho_b), \tag{6}$$

where $\boldsymbol{\theta} = (\varepsilon_a, \varepsilon_b, \varrho_a, \varrho_b, g_a, g_b)$. We allow $\varepsilon_a, \varepsilon_b, g_a, g_b$, and ϱ_a to vary in the range $[0, 2]$, and $\varrho_b \in [0, 1]$. Finally, we vary the structural parameters $a, b \in [0, 5]$, and $\beta \in [1, 5]$.

In Figure 2, we plot the population average effort over time for samples obtained from the marginal likelihood constructed with TBS-SMC, VO-NF with $\gamma = 0.1$ and $\gamma = 0.01$, and a uniform proposal density over $\mathcal{P}$, $q(\boldsymbol{\omega}, \boldsymbol{\theta}) \propto \mathbb{I}[(\boldsymbol{\omega}, \boldsymbol{\theta}) \in \mathcal{P}]$. We furthermore show in Figure 3 contour plots generated by samples from the proposal distribution q constructed with TBS-SMC and VO-NF, in which the marginal distributions are shown on the diagonal and joint bivariate distributions are shown on the lower-diagonal. Samples from TBS-SMC's marginal likelihood are generated with $\epsilon = 0.5$.

From this, we observe that both TBS-SMC and VO-NF have been able to identify regions of $\mathcal{P}$ that produce synthetic populations – and, by extension, scenarios – that satisfactorily match the target scenario as encoded by ℓ in Equation (5): whereas samples from the uniform proposal over $\mathcal{P}$ generates populations and scenarios corresponding to decreasing, increasing, and stagnating levels of effort expended by the labour market participants over time, both TBS-SMC and VO-NF have successfully identified structural and population-level parameters that generate a diverse range of populations and scenarios in which effort levels decrease noticeably over time. Similar behaviour is seen between VO-NF at $\gamma = 0.1$ and TBS-SMC at this simulation budget – for both methods, the range of scenarios generated is reasonably broad, with different degrees of decline observed within the sample scenarios they generate – while VO-NF at $\gamma = 0.01$ gives simulation runs that more strictly apply the selection criteria encoded by ℓ. This exhibits the trade-off between (a) the diversity of synthesised populations and scenarios and (b) the ability to identify the most extreme manifestations of the target scenario discussed in Section 3.

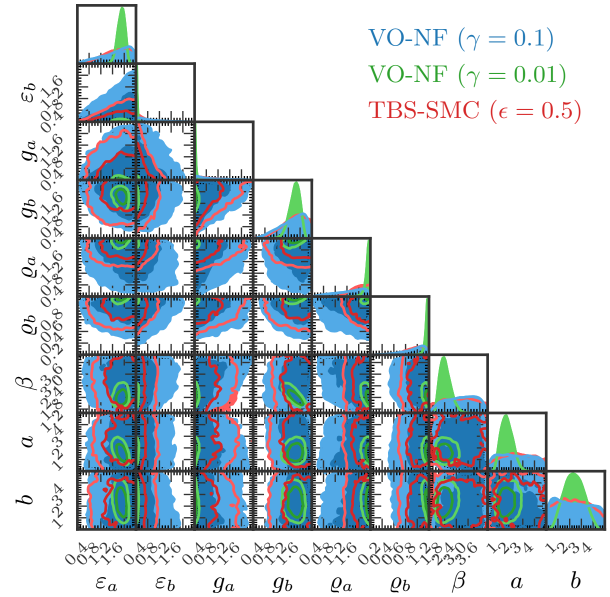

Figure 3: The proposal distributions for (a) VO-NF with $\gamma = 0.1$ (blue) and $\gamma = 0.01$ (green), and (b) TBS-SMC with $\epsilon = 0.5$ (red).

We note that this example demonstrates that TBS-SMC and VO-NF, in conjunction with the use of common parametric distributions to construct $f(\cdot \mid \boldsymbol{\theta})$, can assist the modeller and decision-maker to develop intuition with respect to the population characteristics that give rise to the target scenario. By inspection of Figure 3, we observe that the target scenario is best generated when the agents in the synthesised populations tend to:

- begin with relatively large values of initial effort levels. This is manifested as relatively high density assigned to larger and lower values of ε_a and ε_b, respectively, which translates to a right-skewed Beta distribution over e_n^0;
- exhibit a strong preference for leisure over income. This is manifested as relatively high density assigned to lower and larger values of g_a and g_b, respectively, which translates to a left-skewed Beta distribution over ν_n and, consequently, synthesised populations of agents who are averse to expending effort and who prefer leisure over income;
- reevaluate their positions in the labour market on a relatively frequent basis. This is manifested as relatively high density assigned to large values for ϱ_a and ϱ_b, which has the effect of increasing the mass assigned by the Gamma distribution to higher values for ρ_n, permitting agents to adjust their effort expenditure frequently enough to realise the targeted decline in average population effort.

Taken together, these features of the synthesised population permit the modeller and decision-maker to explain the target scenario in terms of the population characteristics.

4.2 Binary opinion dynamics

We next consider a model of binary opinion dynamics in a population of N agents. At time $t = 0, 1, \ldots, T$, each agent $n = 1, \ldots, N$

in this population holds an opinion $x_n^t \in \{0,1\}$ and is embedded in an undirected graph $\mathcal{G} = (\{1,\dots,N\},\mathcal{E})$, where the set of agents that neighbour n is denoted $\mathcal{N}_n = \{m \mid (n,m) \in \mathcal{E}\}$. The initial opinion x_n^0 of agent $n = 1,\dots,N$ is drawn $x_n^0 \sim \text{Bernoulli}(r)$, where $r \in (0,1)$. At each time step $t \geq 1$, the opinion of agent n updates as $x_n^t \sim \text{Bernoulli}(\alpha(\{x_n^{t-1}\}_{n=1,\dots,N}))$, where

$$\alpha(\{x_n^{t-1}\}_{n=1,\dots,N}) = \begin{cases} \frac{\mu_n}{|\mathcal{N}_n|}\sum_{m\in\mathcal{N}_n} x_m^{t-1}, & x_n^{t-1}=0; \\ 1-\nu_n\left(1-\frac{1}{|\mathcal{N}_n|}\sum_{m\in\mathcal{N}_n} x_m^{t-1}\right), & x_n^{t-1}=1. \end{cases}$$

In the above, $\mu_n \in [0,1]$ (resp. $\nu_n \in [0,1]$) is a parameter capturing the readiness with which agent n transitions from opinion 0 to 1 (resp. 1 to 0) as a result of social pressure from its neighbours $\mathcal{N}_n$. Each agent is thus equipped with three attributes: their initial opinion x_n^0, and the parameters μ_n and ν_n that determine the readiness with which agent n changes, respectively, from opinion 0 to 1 and 1 to 0 as a result of social pressure from its neighbours. This gives $\mathbf{z}_n = (x_n^0, \mu_n, \nu_n) \in \{0,1\}\times[0,1]^2$ as the set of attributes that must be generated in the population synthesis procedure. We simulate this model on graphs generated from the Barabasi-Albert random graph model of preferential attachment [3]; further details are provided in the supplement [16]. For this model, there are no further structural parameters that we are required to tune.

4.2.1 What are the most complex opinion dynamics that can be realised? In this experiment, we aim to generate scenarios in this experiment that correspond to "complex" opinion dynamics within the simulated population, according to some measure of complexity. We consider such scenarios for two main reasons:

(1) In many planning contexts, planners and decision-makers can be interested in guarding against volatility and unpredictability in the environment within which they will make decisions. Simulating such scenarios can therefore be important for testing strategies and policies that are intended to mitigate against the threats posed by turbulent scenarios;
(2) In many models, there exist absorbing states that lead to unrealistic dynamics. A particularly relevant example related to this experiment is a state of consensus achieved in opinion dynamics models, in which all agents agree and share identical opinions. Such states exist within the simulator considered in this section: for example, the states in which all agents share opinions 1 or 0 are absorbing states, leading to a system state from which no further changes can take place. The ability to automatically synthesise populations that do not lead to such unrealistic dynamics when this is not of interest – and instead to more realistic dynamics corresponding to persistent and changing levels of disagreement within the population – is therefore of interest to opinion dynamics modellers in particular.

To target such scenarios with the synthesised population, we consider the following composite loss function:

$$\ell(\mathbf{x}) = \ell_v(\mathbf{x}) + \ell_a(\mathbf{x}).$$

The component

$$\ell_v(\mathbf{x}) = T - \sum_{t=1}^{T}\left\| \frac{1}{N}\sum_{n=1}^{N}\left(x_n^t - x_n^{t-1}\right)\right\|$$

is the negative of the one-variation of the time series of the average opinion within the population, offset by the maximum possible value of T for the one-variation of a piecewise linear curve in $[0,1]$ on the grid $\{0,1,\dots,T\}$. This quantity is often used, informally speaking, to characterise the variability or volatility of a curve. The second component is taken to be

$$\ell_a(\mathbf{x}) = \sum_{t=0}^{T}\left(\prod_{n=1}^{N}\mathbb{I}\left[x_n^t = 0\right] + \prod_{n=1}^{N}\mathbb{I}\left[x_n^t = 1\right]\right),$$

which penalises synthesised populations that lead to time points spent in the two absorbing states discussed previously (i.e., those consisting of all agents possessing opinion 0 or opinion 1). While other loss functions based on, for example, the Hurst exponent [21] or fractal dimension [5] estimated from the time series would also be appropriate, we assume this loss function for simplicity. We take $f(\mathcal{A}_N \mid \boldsymbol{\theta}) = \prod_{n=1}^{N}\iota(\mathbf{z}_n \mid \boldsymbol{\theta})$ with

$$\iota(\mathbf{z}_n \mid \boldsymbol{\theta}) = \text{Bernoulli}\left(x_n^0 \mid r\right)\cdot\text{Beta}\left(\mu_n \mid \varrho_a,\varrho_b\right) \cdot \text{Beta}(\nu_n \mid g_a, g_b), \tag{7}$$

where $\boldsymbol{\theta} = (r,\varrho_a,\varrho_b,g_a,g_b) \in [0,1]\times[0,5]^4$.

In Figure 4, we show the average population opinion observed in samples generated form the marginal likelihood function using TBS-SMC, VO-NF, and a uniform proposal distribution, $q(\boldsymbol{\theta}) \propto \mathbb{I}[\boldsymbol{\theta} \in \Theta]$, while in Figure 5 we show contour plots generated by samples from each of these proposal distributions. The proposal distribution generated by TBS-SMC corresponds to a threshold value of $\epsilon = 400$. As in Section 4.1, we observe that taking a relatively large value of $\gamma = 100$ results in comparable behaviour between VO-NF and TBS-SMC at this simulation budget, both of which generate higher-variance proposal distributions in Figure 5 that translate into a relatively broad variety of scenarios in which the synthesised populations give rise to the complex, changeable dynamics selected by ℓ. In contrast, the comparatively low value of $\gamma = 1$ results once again in a lower-variance proposal distribution, and by extension less diversity in the synthesised populations and scenarios generated. In each of these cases, however, the synthesised populations generate scenarios that resemble the target scenario far more accurately than baseline performance achieved by simply taking a uniform proposal over Θ, which we observe can frequently give rise to trivial dynamics in which the system enters the absorbing state corresponding to complete consensus within the population.

Furthermore, as in the previous experiment, we see that TBS-SMC and VO-NF used in conjunction with the use of common parametric distributions to construct $f(\cdot \mid \boldsymbol{\theta})$ facilitates intuition-building. By inspecting the proposal distributions generated by TBS-SMC and VO-NF with $\gamma = 100$, we see that the scenarios we seek to generate – those exhibiting "complex" dynamics in the averaged population opinion – are more readily generated when agents' opinions are relatively "sticky": low (resp. high) values of ϱ_a and g_a (resp. ϱ_b and g_b) are assigned comparatively high mass by the proposal distributions, and the bivariate joint densities for (ϱ_a,ϱ_b) and (g_a,g_b) show that $\varrho_b > \varrho_a$ and $g_b > g_a$ is generally desirable. Both of these features correspond to Beta distributions over μ_n and ν_n that are left-skewed, corresponding to a synthesised population comprised by more stubborn and less malleable agents.

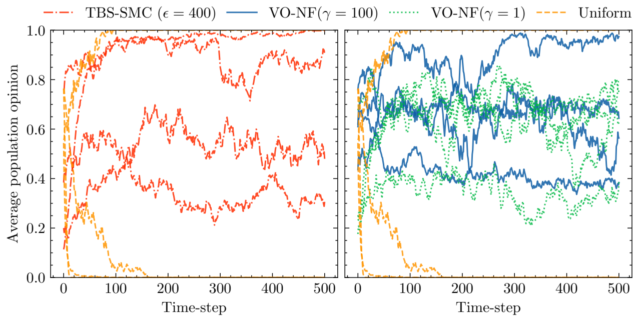

Figure 4: Simulations of the average opinion over time generated from the marginal likelihood function (Equation (1)) using TBS-SMC (red, dot-dashed), VO-NF (blue solid, green dotted), and a Uniform distribution over Θ (orange dashed).

Finally, this experiment highlights the potentially desirable behaviour induced by choosing a far-from-minimal value for ϵ or γ in cases where the loss function ℓ does not perfectly reflect the scenarios of interest to the experimenter. By visual inspection of Figure 4, the sample scenarios generated by VO-NF at $\gamma = 1$ – corresponding to a smaller entropy regularisation term in Equation (4), and thus a narrower proposal q on regions of $\mathbf{\Theta}$ that minimise ℓ – appear *less* complex than those generated by VO-NF at $\gamma = 100$: the former sample paths concentrate slightly more strongly around, and fluctuate less significantly from, an average population opinion of approximately 0.5 (i.e., far from the top and bottom boundaries of the y-axis), while the latter sample paths are less constrained to this region. This suggests that while the choice of ℓ we employ in this experiment does contain information on the "complexity" of the time series, it is misspecified to some degree. In such cases, and as demonstrated by this experiment, it may therefore be desirable to use larger values of γ or ϵ in the VO and TBS approaches we consider, since this can ensure the variety of populations and scenarios is sufficiently diverse, resulting in a higher probability of covering the scenarios of interest and that ℓ imperfectly specifies.

5 RELATED WORK

A closely related body of work revolves around the use of the *Patient Rule Induction Method* (PRIM, see [19, 26]) for scenario discovery and exploratory modelling in simulation models. PRIM identifies rectangular regions in the input space in which the (average) output of the modelled system is an outlier with respect to some baseline behaviour, and has been used in a variety of settings for scenario discovery in simulation models of complex systems. PRIM is related to the TBS approach we consider: in both cases, samples are generated across the parameter space, a subset of which is given further consideration if they generate simulations that produce output that falls beyond some (implicitly or explicitly defined) threshold value. Our work differs from the PRIM methodology in two main respects, by (i) seeking to develop the relationship between scenario generation and population synthesis in ABMs, and (ii) proposing and considering multiple approaches that do not restrict the space of solutions to the space of rectangles on the input space, instead permitting distributions of arbitrary shape and complexity.

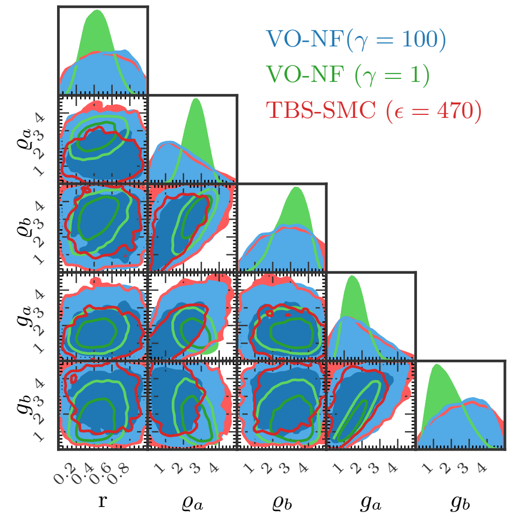

Figure 5: The proposal distribution for (a) VO-NF with $\gamma = 100$ (blue), and $\gamma = 1$ (green), and (b) TBS-SMC with $\epsilon = 400$ (red).

A related problem to that considered in the current work is inverse design in materials science [23, 29, 31], in which a space of possible materials is searched for a material exhibiting a prescribed set of desired behaviours and properties. A number of techniques have been considered to address this problem, including the use of various methods inspired by or drawn from the probabilistic machine learning and generative modelling literature [31]. Our work applies similar ideas in the context of targeted scenario generation in ABMs through appropriately synthesised populations.

Beyond this, there exists a variety of work on constructing realistic synthetic populations for ABMs from data collected on real-world populations [9, 10, 34, 35]. The primary focus of this body of work is to construct synthetic populations that match the distribution of attributes observed in a sample from the real world, for example through iterative proportional fitting [see e.g., 11, 18, 39, 41] or deep generative modelling [9]. One of the main technical challenges here is to account for the under-specified nature of the collected data: such (often census) data is typically only partially revealed to researchers by only specifying certain pairwise relationships between attributes, rather than all relationships between all variables. Furthermore, discreteness is typically introduced artificially through the binning of continuous variables. Our work differs from this body of work, however, in that our approach is designed with the intention of performing exploratory modelling, scenario discovery, and simulation-based planning [see, e.g., 27] with, for example, a (partially) calibrated ABM for which uncertainty regarding the population composition remains.

The general problem of synthesising populations can naturally be related to the problem of synthetic data generation [24, 43]. Our work additionally develops this body of research: in Jordon et al. [Section 3.2, 24], the authors discuss the potential of synthetic data to be employed in what-if scenario generation settings in causal modelling, which aligns closely with the content of our work. In particular, we consider in this work how synthetic data for an artificial population may be generated "in-the-loop" according to the scenario to be generated within the simulated environment.

Finally, the techniques proposed in Section 3 are closely related to simulation-based techniques used in the literature on approximate inference [12, 14, 15, 42]. This relationship is a secondary benefit of our approach to synthesising populations: framing the problem of scenario generation and synthesising populations in ABMs as inference on population-level parameters relates the literature on synthesising populations for ABMs on the one hand, and approximate inference methods for implicit probabilistic models on the other, enabling each of these components of ABMing theory and practice to benefit from the problems and advances seen in the other. In particular, Equation (1) may be seen as a variational program as described in [40], while the methods described in Sections 3.1 and 3.2 can be seen as, respectively, approximate Bayesian computation [4] and (generalised) variational inference [17, 25] for population level parameters in which (a) the loss function ℓ employed within the scenario generation procedure is not restricted to some measure of incongruence between the model behaviour and some observed data, and (b) the employed prior density over $\mathcal{P}$ is uniform (and potentially improper, when $\mathcal{P}$ is unbounded).

6 DISCUSSION

In this paper, we propose and evaluate two broad approaches to population synthesis in agent-based models with the goal of generating and exploring scenarios of interest to an experimenter. This will facilitate the use of agent-based simulators to plan for possible future behaviours of a complex system under uncertainty about the characteristics of the population comprising the system. We demonstrate with simulations that our approaches can accurately identify and generate a diverse range of synthetic populations and, consequently, scenarios of interest through their simulated behaviours.

The approaches we present rely on the experimenter's ability to select an appropriate loss function through which suitable populations and scenarios are generated. What constitutes an appropriate loss functions will in general depend on both the specific model under consideration and the targeted scenario. It is therefore challenging to provide overarching guidance on its design, and we note that this may be a factor which makes the application of our proposed approaches nontrivial in practice.

Our approaches furthermore rely on the identification of a suitable attribute distribution $f(\cdot \mid \boldsymbol{\theta})$. In many cases, this will simply be a modelling choice and entirely for the modeller to choose. There nonetheless exists multiple possible strategies that can be taken to construct f. The strategy adopted in our simulation studies is to construct f as a factorised distribution consisting of standard distributions over each agent attribute. Each factor is then tailored to the specific domain of the corresponding attribute, e.g., the use of Beta distributions for attributes lying in the unit range. However, our approaches are not restricted to the use of simple parametric or factorised distributions alone: greater expressivity can be achieved (possibly at the expense of interpretability, however) using, e.g., mixture models [6], or other explicit or implicit generative models (e.g., normalising flows [37], GANs [20], or deep graph models [28]).

The question of which of the two approaches discussed in this article, threshold-based sampling or variational optimisation, may be most useful to the practitioner arises naturally. We generally observed VO-NF to converge within a few thousand simulations across different values of γ, whereas many more simulations were required to construct the TBS proposal distribution. This suggests that VO may generally be preferable. However, since γ indirectly controls the variance of the marginal likelihood and generated values for ℓ by directly controlling the entropy of the proposal, it can be harder to calibrate γ to the desired degree of congruence between the simulated scenarios and the target scenario. In contrast, ϵ is likely to be more comprehensible to the experimenter – if the experimenter has sufficient intuition to designate an appropriate loss function ℓ for the scenarios they wish to target, it seems reasonable to expect that this intuition will extend to identifying minimally desirable distances ϵ from it – meaning that the "right" amount of diversity in the synthesised populations and scenarios may be achieved more easily with TBS. For this reason, TBS may therefore be more easily applied than VO, despite the larger simulation burden we observed when constructing a single proposal distribution. Surrogate models and alternative (e.g., quasi-Monte Carlo [13]) samplers can also be employed to reduce the computational burden. Further work will nonetheless be required to more comprehensively explore the relative merits of TBS and VO.

Although our discussion has refrained from incorporating prior beliefs, each of the approaches we consider can be immediately incorporated into a Bayesian framework by applying pointwise weightings to the proposal distribution using a further distribution $\pi : \mathcal{P} \rightarrow \mathbb{R}_{\geq 0}$. For example, this might be a prior distribution used to encode the decision-maker's *a priori* beliefs about (un)likely values for $(\boldsymbol{\omega}, \boldsymbol{\theta})$, or a posterior distribution obtained by performing Bayesian inference on $(\boldsymbol{\omega}, \boldsymbol{\theta})$ using any data that is available. In such cases, the approaches discussed in Sections 3.1 and 3.2 are closely related to approximate Bayesian computation (ABC, [4]) and (generalised) variational Bayesian inference (GVI, [17, 25]). However, despite this similarity, our motivation deviates significantly from the typical use-cases for ABC and GVI: ABC and GVI are typically employed when the experimenter is concerned with drawing inferences about parameter values upon observing a dataset from a real-world system being modelled by the simulation model; the problem that we consider however differs in the sense that data is not used to draw inferences about parameter values, but instead parameter values are proposed to generate desired model behaviour without relying on the presence of data and with the purpose of planning for scenarios that are of interest to the experimenter.

ACKNOWLEDGMENTS

This research was supported by a UKRI AI World Leading Researcher Fellowship awarded to Wooldridge (grant EP/W002949/1). M. Wooldridge and A. Calinescu 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. J. D. Farmer acknowledges funding from Baillie Gifford.

REFERENCES

[1] Robert Axtell. 1999. *The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions.*

[2] Robert L. Axtell and J. Doyne Farmer. [n.d.]. Agent-Based Modeling in Economics and Finance: Past, Present, and Future. *Journal of Economic Literature* ([n. d.]). https://doi.org/10.1257/jel.20221319

[3] Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *science* 286, 5439 (1999), 509–512.

[4] Mark A Beaumont, Jean-Marie Cornuet, Jean-Michel Marin, and Christian P Robert. 2009. Adaptive Approximate Bayesian Computation. *Biometrika* 96, 4 (2009), 983–990.

[5] Benoit Mandelbrot. 1967. How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. *Science (New York, N.Y.)* 156, 3775 (1967), 636–638. https://doi.org/10.1126/science.156.3775.636 arXiv:https://www.science.org/doi/pdf/10.1126/science.156.3775.636

[6] Christopher M. Bishop. 1994. Mixture Density Networks. *Mixture density networks* (1994).

[7] David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. 2017. Variational Inference: A Review for Statisticians. *J. Amer. Statist. Assoc.* 112, 518 (April 2017), 859–877. https://doi.org/10.1080/01621459.2017.1285773 arXiv:1601.00670 [cs, stat]

[8] Eric Bonabeau. 2002. Agent-Based Modeling: Methods and Techniques for Simulating Human Systems. *Proceedings of the National Academy of Sciences of the United States of America* 99, 10 (2002), 7280–7287. arXiv:3057854

[9] Stanislav S. Borysov, Jeppe Rich, and Francisco C. Pereira. 2019. Scalable Population Synthesis with Deep Generative Modeling. *Transportation Research Part C: Emerging Technologies* 106 (Sept. 2019), 73–97. https://doi.org/10.1016/j.trc.2019.07.006 arXiv:1808.06910 [cs, stat]

[10] Kevin Chapuis, Patrick Taillandier, and Alexis Drogoul. 2022. Generation of Synthetic Populations in Social Simulations: A Review of Methods and Practices. *Journal of Artificial Societies and Social Simulation* 25, 2 (2022), 6.

[11] Abdoul-Ahad Choupani and Amir Reza Mamdoohi. 2016. Population Synthesis Using Iterative Proportional Fitting (IPF): A Review and Future Research. *Transportation Research Procedia* 17 (2016), 223–233. https://doi.org/10.1016/j.trpro.2016.11.078 International Conference on Transportation Planning and Implementation Methodologies for Developing Countries (12th TPMDC) Selected Proceedings, IIT Bombay, Mumbai, India, 10-12 December 2014.

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

[13] Josef Dick, Frances Y Kuo, and Ian H Sloan. 2013. High-dimensional integration: the quasi-Monte Carlo way. *Acta Numerica* 22 (2013), 133–288.

[14] Joel Dyer, Patrick Cannon, J. Doyne Farmer, and Sebastian M. Schmon. 2024. Black-box Bayesian inference for agent-based models. *Journal of Economic Dynamics and Control* (2024).

[15] Joel Dyer, Patrick W Cannon, and Sebastian M Schmon. 2022. Amortised likelihood-free inference for expensive time-series simulators with signatured ratio estimation. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 11131–11144.

[16] Joel Dyer and Arnau Quera-Bofarull. 2024. *joelnmdyer/synthpop: AAMAS release.* https://doi.org/10.5281/zenodo.10629106

[17] Joel Dyer, Arnau Quera-Bofarull, Ayush Chopra, J Doyne Farmer, Anisoara Calinescu, and Michael Wooldridge. 2023. Gradient-assisted calibration for financial agent-based models. In *Proceedings of the Fourth ACM International Conference on AI in Finance*. 288–296.

[18] W. Edwards Deming Frederick F. Stephan and Morris H. Hansen. 1940. The Sampling Procedure of the 1940 Population Census. *J. Amer. Statist. Assoc.* 35, 212 (1940), 615–630. https://doi.org/10.1080/01621459.1940.10502056 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1940.10502056

[19] Jerome H Friedman and Nicholas I Fisher. 1999. Bump hunting in high-dimensional data. *Statistics and computing* 9, 2 (1999), 123–143.

[20] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Networks. https://doi.org/10.48550/arXiv.1406.2661 arXiv:1406.2661 [cs, stat]

[21] H. E. Hurst. 1951. Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers* 116, 1 (1951), 770–799. https://doi.org/10.1061/TACEAT.0006518 arXiv:https://ascelibrary.org/doi/pdf/10.1061/TACEAT.0006518

[22] W. K. Hastings. 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika* 57, 1 (1970), 97–109. https://doi.org/10.2307/2334940 arXiv:2334940

[23] Jiaqi Jiang and Jonathan A Fan. 2019. Global optimization of dielectric metasurfaces using a physics-driven neural network. *Nano letters* 19, 8 (2019), 5366–5372.

[24] James Jordon, Lukasz Szpruch, Florimond Houssiau, Mirko Bottarelli, Giovanni Cherubin, Carsten Maple, Samuel N Cohen, and Adrian Weller. 2022. Synthetic Data–What, Why and How? *arXiv preprint arXiv:2205.03257* (2022). arXiv:2205.03257

[25] Jeremias Knoblauch, Jack Jewson, and Theodoros Damoulas. 2022. An Optimization-Centric View on Bayes' Rule: Reviewing and Generalizing Variational Inference. *The Journal of Machine Learning Research* 23, 1 (2022), 5789–5897.

[26] Jan H. Kwakkel and Marc Jaxa-Rozen. 2016. Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes. *Environmental Modelling & Software* 79 (2016), 311–321. https://doi.org/10.1016/j.envsoft.2015.11.020

[27] Jan H Kwakkel and Erik Pruyt. 2013. Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change* 80, 3 (2013), 419–431.

[28] Yujia Li, Oriol Vinyals, Chris Dyer, Razvan Pascanu, and Peter Battaglia. 2018. Learning deep generative models of graphs. *arXiv preprint arXiv:1803.03324* (2018).

[29] Dianjing Liu, Yixuan Tan, Erfan Khoram, and Zongfu Yu. 2018. Training deep neural networks for the inverse design of nanophotonic structures. *Acs Photonics* 5, 4 (2018), 1365–1369.

[30] Jun S. Liu and Rong Chen. 1998. Sequential Monte Carlo Methods for Dynamic Systems. *J. Amer. Statist. Assoc.* 93, 443 (1998), 1032–1044. https://doi.org/10.2307/2669847 arXiv:2669847

[31] Wei Ma, Feng Cheng, Yihao Xu, Qinlong Wen, and Yongmin Liu. 2019. Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy. *Advanced Materials* 31, 35 (2019), 1901111.

[32] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 2004. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* 21, 6 (Dec. 2004), 1087–1092. https://doi.org/10.1063/1.1699114

[33] Shakir Mohamed, Mihaela Rosca, Michael Figurnov, and Andriy Mnih. 2020. Monte Carlo Gradient Estimation in Machine Learning. *Journal of Machine Learning Research* 21, 132 (2020), 1–62.

[34] Kirill Müller and Kay W. Axhausen. 2010. Population Synthesis for Microsimulation: State of the Art. *Arbeitsberichte Verkehrs- und Raumplanung* 638 (Aug. 2010). https://doi.org/10.3929/ethz-a-006127782

[35] Tadahiko Murata, Daiki Iwase, and Takuya Harada. 2022. Workplace assignment to workers in synthetic populations in Japan. *IEEE Transactions on Computational Social Systems* (2022).

[36] Kerri-Ann Norton, Chang Gong, Samira Jamalian, and Aleksander S. Popel. 2019. Multiscale Agent-Based and Hybrid Modeling of the Tumor Immune Microenvironment. *Processes (Basel, Switzerland)* 7, 1 (Jan. 2019), 37. https://doi.org/10.3390/pr7010037

[37] George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. 2021. Normalizing Flows for Probabilistic Modeling and Inference. *Journal of Machine Learning Research* 22, 1, Article 57 (Jan. 2021).

[38] George Papamakarios, Theo Pavlakou, and Iain Murray. 2017. Masked Autoregressive Flow for Density Estimation. In *Advances in Neural Information Processing Systems*, Vol. 30. Curran Associates, Inc.

[39] David R Pritchard and Eric J Miller. 2012. Advances in population synthesis: fitting many attributes per agent and fitting to household and person margins simultaneously. *Transportation* 39, 3 (2012), 685–704.

[40] Rajesh Ranganath, Dustin Tran, Jaan Altosaar, and David Blei. 2016. Operator Variational Inference. In *Advances in Neural Information Processing Systems*, D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Eds.), Vol. 29. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2016/file/d947bf06a885db0d477d707121934ff8-Paper.pdf

[41] Ludger Ruschendorf. 1995. Convergence of the Iterative Proportional Fitting Procedure. *The Annals of Statistics* 23, 4 (1995), 1160–1174. http://www.jstor.org/stable/2242759

[42] Alvaro Tejero-Cantero, Jan Boelts, Michael Deistler, Jan-Matthis Lueckmann, Conor Durkan, Pedro J. Gonçalves, David S. Greenberg, and Jakob H. Macke. 2020. Sbi: A Toolkit for Simulation-Based Inference. *Journal of Open Source Software* 5, 52 (Aug. 2020), 2505. https://doi.org/10.21105/joss.02505

[43] Boris van Breugel and Mihaela van der Schaar. 2023. Beyond Privacy: Navigating the Opportunities and Challenges of Synthetic Data. *arXiv preprint arXiv:2304.03722* (2023).

[44] John von Neumann. 1951. Various Techniques Used in Connection with Random Digits. In *Monte Carlo Method*, A.S. Householder, G.E. Forsythe, and H.H. Germond (Eds.). National Bureau of Standards Applied Mathematics Series, 12, Washington, D.C.: U.S. Government Printing Office, 36–38.

[45] Xin Ye, Karthik Konduri, Ram Pendyala, Bhargava Sana, and Paul Waddell. 2009. Methodology to Match Distributions of Both Household and Person Attributes in Generation of Synthetic Populations. (Jan. 2009).