Equation-free analysis of agent-based models: systematic parameter determination

Spencer A. Thomas[1], David J.B. Lloyd[1], and Anne C. Skeldon[1]

[1]Department of Mathematics, University of Surrey, Guildford, GU2 7XH, UK

April 21, 2015

1 Time horizon τ

For deterministic systems, an ensemble of microscopic models initialised at the same state, will all be in the same state after a simulation time of τ. Stochastic systems however, will result in distribution of final states with a mean $\bar{x}$ and variance $\sigma^2 \neq 0$. The value of σ^2 will vary due to the dynamics and the noise of the system, as may $\bar{x}$, and thus depends on the size of the time evolution. For increasing time windows, σ^2 may either;

1. increase indefinitely with τ (diffusion in an open system),
2. increase to a point then remain approximately constant (point of balance or symmetry in the system), or
3. increase to a maximum point then begin to reduce (transition from one state to another).

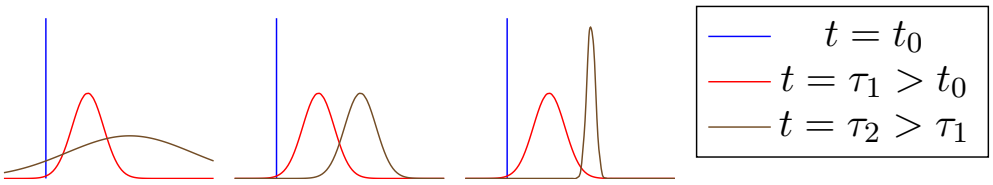

Figure 1: *Dependence of σ^2 on τ. For increasing τ, σ^2 increases; indefinitely (left), to a constant (mid), or to a maximum before reducing (right).*

This is illustrated in Fig. 1. For an arbitrary system, the distribution of an ensemble of simulation, the realisations, at τ is unknown, and although we can simply calculate $\bar{x}$, we can not calculate σ^2 in general. We can however average a number of ensembles resulting in a distribution of $\bar{x}$ which, according to the central limit theorem (CLT) is normally distributed. From the distribution of $\bar{x}$ we can calculate a mean $\overline{\bar{x}}$ (the mean of the means) and a standard deviation ω. As increasing τ leads to an increase, at least initially, in $\bar{x}$ and σ^2, it will also effect $\overline{\bar{x}}$ and ω. By defining a a minimum threshold γ for the value of ω, which reflects the amount of variation across the ensembles, we can obtain a value of τ that enables sufficient time evolution such that the dynamic behaviour in the

model is observable. That is, increase τ until $\omega \geq \gamma$. This provides a platform for the systematic determination of τ for a specific model, avoiding excessive trial and error testing, or the need of a deterministic equivalent. This process is illustrated in Fig. 2 and the procedure is summarised in Algorithm 1.

Algorithm 1 Time horizon predictor.

```
# obtain initial parameter settings
N = N_int # initial realisations (default=10)
τ = τ_int # initial time horizon
δτ # time horizon increment
M # number of runs to average over
MaxIter # maximum number of iterations
γ # threshold for dynamics
while iterations < MaxIter do
    for j = 0; j < M do
        LIFT(N)
        SIMUALTE(N, τ)
        x̄[j] = RESTRICT(N)
        j := j+1
    end for
    # 95 % confidence interval of distribution of means
    ω_i = 95ConfInt(x̄)
    if ω > γ then
        EXIT # τ large enough to see some dynamics
    end if
    τ := τ * δτ
    iterations := iterations + 1
end while
```

2 Number of microscopic realisations N

Once we have obtained a value for the time evolution of the microscopic simulations we can then obtain a prediction for the number of microscopic realisations required to perform EF continuation of a generic ABM. The time horizon predictor in Algorithm 1 provides both a prediction of τ and a measure of how well defined this value is through ω_N, which we can use as our threshold in the realisation predictor. That is, we can again use the CLT to obtain a (Gaussian) distribution of means and calculate the mean of means, variance and confidence interval. Here we can use the value of ω from the time horizon predictor as a limit for the variance of the distribution of means for the realisations, i.e. ω_N. This process is summaries in Algorithm 2.

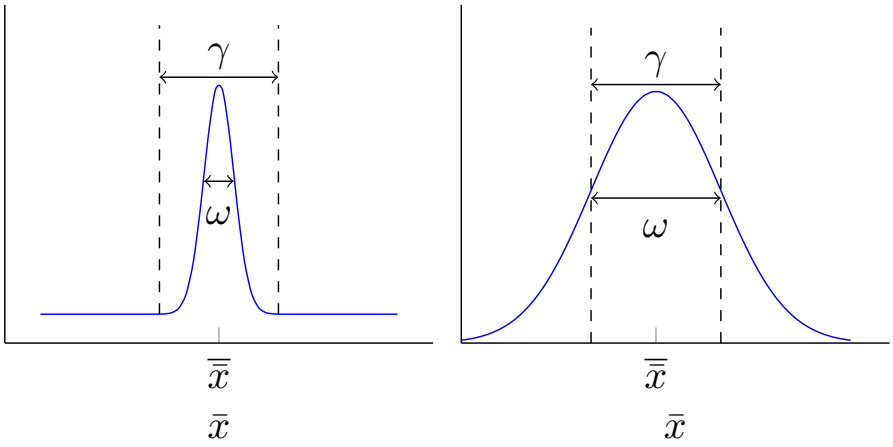

Figure 2: *Variance of the distribution of means, ω with γ limit. For small τ (left) $\omega < \gamma$, whereas for larger τ (right) $\omega \approx \gamma$.*

Algorithm 2 Realisation predictor.

```
# obtain initial parameter settings
N = N_int # initial realisations
δN # increment in N
M # number of runs to average over
N_Max # maximum number of realisations
τ # time horizon from Algorithm 1
γ # threshold for dynamics from Algorithm 1
while ω_N > γ do
    for j = 0; j < M do
        LIFT(N)
        SIMUALTE(N, τ)
        x̄[j] = RESTRICT(N)
        j := j+1
    end for
    # 95 % confidence interval of distribution of means
    ω_N = 95ConfInt(x̄)
    N := N * δN
end while
```

3 Continuation step size δs and finite differencing

We require a step size (δs), in the continuation parameter λ such that if $f(x^*, \lambda) = x^*$, then $f(x^*, \lambda + \delta s) \approx x^*$. That is, we require δs to be small enough so that we obtain a suitable prediction to f under this perturbation. However this cannot be arbitrary small due to finite computational resources, impractical run times, and will depend on the specific problem under investigation. As with N and τ, for δs we can determine a value based directly on simulation of the system. In this case by examining the level for stochasticity through the variance using the parameter obtained with Algorithms 1 and 2. We perform the *Lift* and *Simulation* operations for one set of realisations and calculate the standard error in the mean (Δ) of the distribution of simulation results. We then

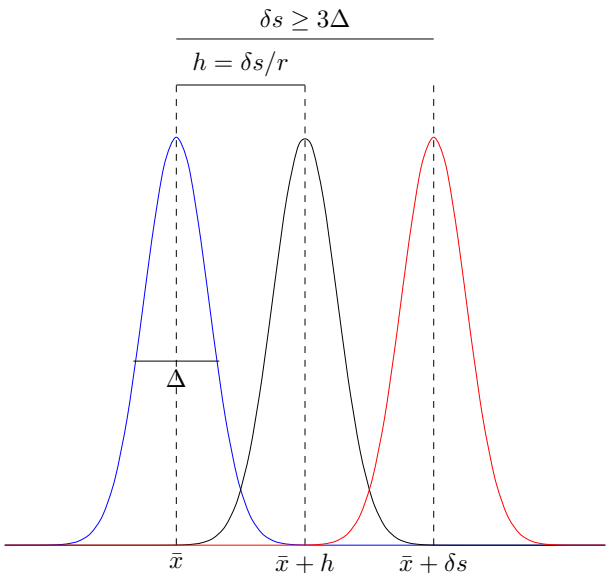

Figure 3: *Illustration of the δs predictor. The standard error (Δ) from the the simulation of N realisations (blue) is used to determine the step size for finite differencing (black) and continuation step size (red).*

simply define the the step size as

$$\delta s = 3\Delta \ . \tag{3.1}$$

During the predictor-corrector continuation used here, it is common to employ a Newton-Rhapson corrector method. In order to determine the update step size in this procedure we use finite differencing to calculate the Jacobian J,

$$J_{ik} = \frac{f_i(x+h) - f_k(x)}{h} \ , \tag{3.2}$$

for $i, k = 0 \ldots m$, where m is the number of orders in f. The result of Eq. (3.2) is an $m \times m$ matrix. The value h is also problem specific as the accuracy of the determination of J will depend on the magnitude of h compared to x. To ensure that J is accurately calculated for a general system, we define the value of h in terms of the continuation step size,

$$h = \frac{\delta s}{r} \ , \tag{3.3}$$

where $r = 2$ in this work.

4 Iterative determination and the selection of γ

It is worth noting that the value of γ in Algorithm 1 and 2 is technically a tunable, problem specific parameter. However, here we have replaced three problem specific parameters with a single parameter, thus reducing the overhead. We note however that γ can easily be obtained for a arbitrary model. By perform the procedure in Algorithm 1 with an arbitrarily large γ (say 10,000) one can monitoring ω as a function of τ. The profile of this curve can be used to select a new value of γ for the specific model and Algorithm 1 can be repeated. Figure 4 shows the dependence of ω on τ when using an arbitrarily large γ. Here we can see that we can use these curves to obtain a better guess for γ and obtain our problem specific equation-free parameters.

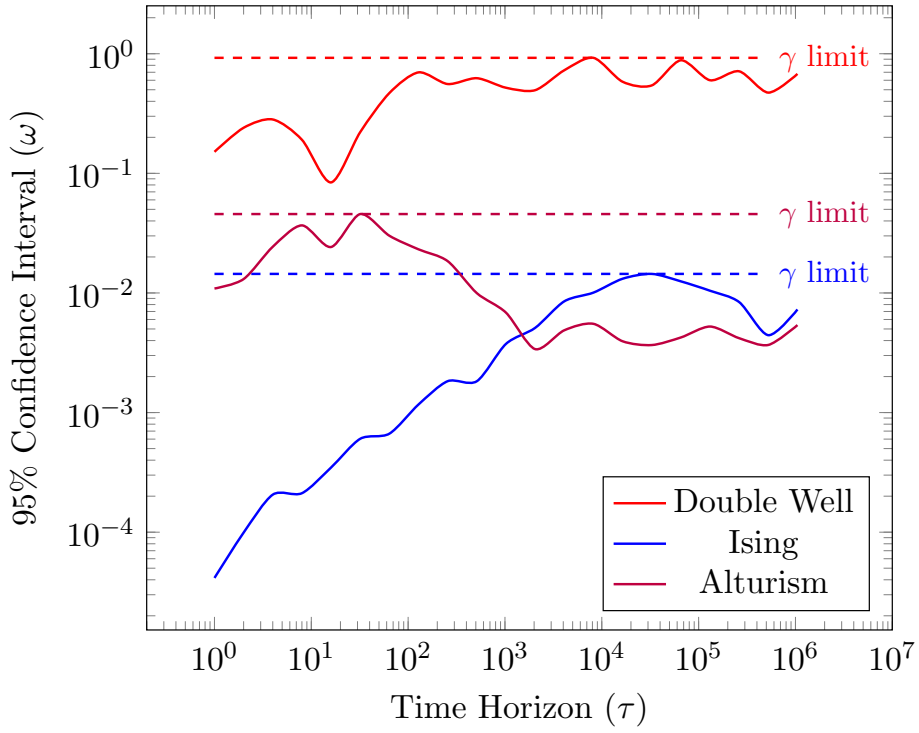

Figure 4: *Confidence interval on mean of $g(x)$ for 3 examples from the Repository of Models showing the maximum value of γ in each case.*

Acknowledgements

The support of the UK Engineering and Physical Sciences Research Council for programme grant EP/H021779/1 (Evolution and Resilience of Industrial Ecosystems (ERIE)) is gratefully acknowledged.