Discussion Paper Series

RIEB

Kobe University

DP2014-09

Perfect Simulation for Models of Industry Dynamics

Takashi KAMIHIGASHI
John STACHURSKI

March 13, 2014

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Perfect Simulation for Models of Industry Dynamics*

Takashi Kamihigashi[a,b] and John Stachurski[c]

March 7, 2014

[a] *IPAG Business School, Paris, France*

[b] *Research Institute for Economics and Business Administration, Kobe University*

[c] *Research School of Economics, Australian National University*

Abstract

In this paper we introduce a technique for perfect simulation from the stationary distribution of a standard model of industry dynamics. The method can be adapted to other, possibly non-monotone, regenerative processes found in industrial organization and other fields of economics. The algorithm we propose is a version of coupling from the past. It is straightforward to implement and exploits the regenerative property of the process in order to achieve rapid coupling.

Keywords: Regeneration, simulation, coupling from the past, perfect sampling

1 Introduction

As is well-known, coupling from the past (CFTP) algorithms can generate perfect samples from otherwise intractable stationary distributions (see, e.g., Propp and Wilson

*We are grateful for financial support from ARC Discovery Outstanding Researcher Award DP120100321 and the Japan Society for the Promotion of Science. Email: `tkamihig@rieb.kobe-u.ac.jp`, `john.stachurski@anu.edu.au`.

(1996)). This makes them an attractive alternative to ordinary forward simulation, where errors are typically difficult to assess. Parallelization is simple in most cases, making it possible to rapidly generate large numbers of independent and identically distributed draws. Because the resulting sample is IID and exact, it can be used to obtain unbiased estimates and confidence intervals for moments and distributions of interest.

The CFTP technique is often used for models with large but discrete state spaces, where standard methods for computing stationary distributions are infeasible.[1] More recently, researchers have developed techniques for implementing CFTP methods in continuous state settings. Murdoch and Green (1998) showed that CFTP can in principle be used in continuous state settings when the underlying Markov process satisfies Doeblin's condition. This condition requires the existence of a nonnegative lower bound function that (a) integrates to a positive value, (b) depends only on the next state, and (c) is pointwise dominated by the transition density function (which depends on both the current state and the next). Theoretical work along the same lines can be found in Foss and Tweedie (1998) and Athreya and Stenflo (2003).

Although these results are fundamental, they can be difficult to apply. Murdoch and Green (1998) admit that their basic method, which is in principle applicable to our model, may have "a limited range of application for two reasons." First, the function associated with Doeblin's condition "may be too small for practical use" to generate exact draws in a reasonable length of time. Second, their method requires the user to draw from scalar multiples of the lower bound transition density and a residual kernel. It can be nontrivial or even impossible to explicitly calculate and draw from these distributions. If approximations are required, this to some degree defeats the purpose of exact sampling.

For these reasons, CFTP methods tend to be popular only in specific settings, perhaps the most notable of which is where the underlying Markov process is stochastically monotone. For such processes, efficient and straightforward CFTP methods are available. Corcoran and Tweedie (2001) developed general results on CFTP particularly suitable for monotone Markov processes. An application to economics can be found in Nishimura and Stachurski (2010), where monotonicity makes the algorithm straightforward to implement and analyze.

In this paper we study regenerative processes associated with industry dynamics that

[1] Applications range from statistical mechanics to page ranking and the design of peer-to-peer file sharing systems. See, for example, Propp and Wilson (1996), Kijima and Matsui (2006), Huber (2003) and Levin et al. (2009).

take the form

$$\Phi_{t+1} = g(\Phi_t, U_{t+1})\mathbb{1}\{\Phi_t \geq x\} + Z_{t+1}\mathbb{1}\{\Phi_t < x\}. \tag{1}$$

Here Φ_t is the state variable taking values in a closed interval, x is a point in the interior of that interval, g is a given function, $\mathbb{1}\{\cdot\}$ is an indicator function and $\{U_t\}$ and $\{Z_t\}$ are IID processes. The function g is monotone increasing in its first argument, and such that the state is driven below x with positive probability in finite steps. When $\Phi_t < x$ the process regenerates, being drawn from the distribution of Z_{t+1}. If this distribution is stochastically dominated by that of $g(x, U_{t+1})$, then the process is stochastically monotone in the usual sense (see, e.g., Corcoran and Tweedie (2001)). As we discuss below, this is not the case in many models of interest. The purpose of this paper is to develop a CFTP algorithm that is designed to produce exact draws from the stationary distribution even in the non-monotone case.

Regenerative models of the form (1) appear in many models of industry dynamics—typical examples include Hopenhayn (1992), Melitz (2003) and Cooley and Quadrini (2001). Regeneration plays an essential role in the theory of industrial organization, where Schumpeter's notion of "creative destruction" summarizes the idea that new and more productive firms replace older ones, rejuvenating overall economic activity (Schumpeter (1942)). In this context, the model in (1) has the following interpretation: A large number of firms produce in a given industry. While incumbent, their productivity evolves according to $\Phi_{t+1} = g(\Phi_t, U_{t+1})$, but when their productivity drops below a threshold level x they exit, and are replaced by a new firm with productivity drawn according to the distribution of Z_t. The value x is obtained by the solution to a dynamic programming problem, taking in to account revenue, costs and so on (although in this paper we take it as given). The stationary distribution of the model can be regarded as representing the time-invariant cross-sectional distribution of productivity in the industry.

The assumption that g is increasing in its first argument is based on the idea that if firm A has higher productivity than firm B today, then firm A is expected to have higher productivity than firm B tomorrow provided that both firms survive. On the other hand, if firm B does not survive and the productivity of firm A is close to the threshold level x, then a new entrant replacing firm B is likely to be more productive than firm A. This characteristic of entry and exit makes the entire process non-monotone.

Other applications for the regenerative model (1) can be found in various intertemporal decision problems. For example, consider the problem of optimal replacement of a part or machine, the performance of which degrades stochastically over time. Typically, the solution to the planning problem involves replacement when some measure of performance falls below a certain threshold (e.g., Rust (1987)). A similar idea is rebooting or restarting operating systems, as discussed by Cotroneo et al. (2014). The performance of these systems can also degrade over time as a result of memory leaks, software conflicts and so on. An essential feature of these models is again non-monotonicity: the very purpose of replacement and rebooting is to break monotonicity and rejuvenate performance.

While we do not assume that the entire process (1) is monotone, we do assume that g is increasing in its first argument, as discussed above, and we heavily exploit this property in our algorithm. We show that the algorithm terminates successfully in finite time with probability one by using both the monotonicity of productivity for incumbents and the regenerative property introduced by new entrants. Our algorithm is distinct from Murdoch and Green's method discussed above (Murdoch and Green (1998)), in that it does not use Doeblin's condition, and does not require explicit knowledge of the transition density.[2] As long as one can simulate the overall Markov process, one can sample exactly from the stationary distribution using the algorithm.[3]

2 Preliminaries

In this section we briefly review a benchmark model of firm dynamics due to Hopenhayn (1992). The model is set in a competitive industry where entry and exit is endogenously determined. In the model there is a large number of firms that produce a homogeneous good. The firms face idiosyncratic productivity shocks that follow a Markov process on $S := [0, 1]$. The conditional cumulative distribution function for the shock process is denoted by $F(\phi' \mid \phi)$. Following Hopenhayn (1992), we impose the following restrictions:

Assumption 2.1. F is decreasing in its second argument and, for any $\epsilon > 0$ and any $\phi \in S$, there exists an integer n such that $F^n(\epsilon \mid \phi) > 0$.[4]

[2] The assumptions used to show the probability one termination of the algorithm in fact imply Doeblin's condition for some n-step transition, but our proof of this property does not use the latter.

[3] As usual, exactness is modulo the errors associated with floating point arithmetic, which cannot be avoided.

[4] $F^n(\cdot \mid \phi)$ is the conditional distribution for productivity after n periods, given current productivity ϕ.

We let P denote the corresponding stochastic kernel. That is, $P(\phi, A) := \int_A F(d\phi' \,|\, \phi)$ for $\phi \in S$ and $A \in \mathscr{B}$, where $\mathscr{B}$ represents the Borel sets on $[0,1]$. Incumbent firms exit the industry whenever their current productivity falls below a reservation value x_t. Letting M_t be the mass of entrants at time t and ν be the Borel probability measure from which the productivity of entrants is drawn, the sequence of firm distributions $\{\mu_t\}$ on S satisfies $\mu_{t+1}(A) = \int P(\phi, A)\mathbb{1}\{\phi \geq x_t\}\mu_t(d\phi) + M_{t+1}\nu(A)$ for all $A \in \mathscr{B}$. At the stationary equilibrium, both x and M are constant, and a stationary distribution μ is a Borel probability[5] measure μ satisfying

$$\mu(A) = \int P(\phi, A)\mathbb{1}\{\phi \geq x\}\mu(d\phi) + M\nu(A) \qquad (A \in \mathscr{B}). \tag{2}$$

It follows from (2) and $\mu(S) = P(\phi, S) = \nu(S) = 1$ that $M = M(x, \mu) := \mu\{\phi \in S : \phi < x\}$. As a result, we can also write (2) as

$$\mu(A) = \int Q(\phi, A)\mu(d\phi) \tag{3}$$

where

$$Q(\phi, A) := P(\phi, A)\mathbb{1}\{\phi \geq x\} + \nu(A)\mathbb{1}\{\phi < x\}. \tag{4}$$

Equation (3) states that μ is a stationary distribution for the stochastic kernel Q in the usual sense of time invariance. As shown by Hopenhayn (1992), the kernel Q has only one stationary distribution. For the purposes of this paper we will treat x as given. For typical parameter values the stationary distribution has no analytical solution.

It is straightforward to produce an ergodic Markov process suitable for simulation such that its stationary distribution coincides with the distribution μ in (3). In essence, we need a method for sampling from the stochastic kernel Q. The first step is to simulate from the conditional distribution $P(\phi, \cdot) = F(\cdot \,|\, \phi)$. In particular, we seek a random variable U and a function g such that $\mathcal{D}(g(\phi, U)) = F(\cdot \,|\, \phi)$ for all $\phi \in S$. (Here $\mathcal{D}(X)$ indicates the distribution of random variable X.) This can be achieved via the inverse transform method, where U is uniform on $[0,1]$ and $g(\phi, u) = F^{-1}(u \,|\, \phi)$.[6] Now consider the process $\{\Phi_t\}$ defined by

$$\Phi_{t+1} = g(\Phi_t, U_{t+1})\mathbb{1}\{\Phi_t \geq x\} + Z_{t+1}\mathbb{1}\{\Phi_t < x\} \tag{5}$$

where $\{(U_t, Z_t)\}$ is IID with $\mathcal{D}(Z_t) = \nu$ and $\mathcal{D}(U_t) = \text{Uniform}[0,1]$. Comparing (4) and (5), it can be seen that $\{\Phi_t\}$ is a Markov process with stochastic kernel Q.

[5] We focus only on normalized measures, since other cases are just scalar multiples.

[6] Here $F^{-1}(\cdot \,|\, \phi)$ is the generalized inverse of $F(\cdot|\phi)$. That is, $F^{-1}(u \,|\, \phi) := \inf\{z : F(z \,|\, \phi) \geq u\}$.

3 Exact Sampling

Let $\{(U_t, Z_t)\}_{t\in\mathbb{Z}}$ be an infinite sequence of IID shocks indexed on $\mathbb{Z}$ and with each pair (U_t, Z_t) having the product distribution Uniform$[0,1] \times \nu$. To simplify notation we will let $g_t := g(\cdot, U_t)$, so that, for example, $g_t \cdots g_1 \, \phi := g_t \circ g_{t-1} \circ \cdots \circ g_1(\phi)$ is exogenous productivity at t, given time zero productivity $\phi \in S$. To further simplify notation, let

$$h_t(\phi) := g(\phi, U_t)\mathbb{1}\{\phi \geq x\} + Z_t \mathbb{1}\{\phi < x\},$$

so that (5) becomes $\Phi_{t+1} = h_{t+1}\, \Phi_t := h_{t+1}(\Phi_t)$.

Now fix $T \geq 1$. For each $\phi \in S$, there is a corresponding "tracking process" that starts at time $-T$ with value ϕ, and then updates with maps $h_{-T+1}, h_{-T+2}, \ldots, h_0$, obtaining the value $h_0 \cdots h_{-T+1}\, \phi$ at time zero. We say that the tracking processes *coalesce* if, for some $T \in \mathbb{N}$, the set of final values

$$h_0 \cdots h_{-T+1}(S) := \{h_0 \cdots h_{-T+1}\, \phi \;:\; \phi \in S\} \tag{6}$$

is a singleton. What we will now show is that under mild conditions coalescence occurs with probability one, and, moreover, that it is not necessary to keep track of the full continuum of tracking processes in order to find the value of the singleton. In particular, we show that, conditional on a certain event described below, the set of final values $h_0 \cdots h_{-T+1}(S)$ has only *finitely* many possibilities. Hence coalescence occurs whenever these finite possibilities take the same value. All of these finite possibilities are computable. To begin describing them, let $T > 1$ be given, let

$$\Sigma_T := \{k \in \mathbb{N} : 1 \leq k < T \text{ and } g_{-T+k} \cdots g_{-T+2} \cdot g_{-T+1} 1 < x\},$$

and let $\sigma_T := \min \Sigma_T$. Intuitively, σ_T is the number of periods that an incumbent firm survives, given that it starts at time $-T$ with maximal productivity 1 and faces the shock sequence $\{U_t\}_{t\in\mathbb{Z}}$. Clearly σ_T is only defined when Σ_T is nonempty. However, the probability that Σ_T is nonempty converges to one as $T \to \infty$ by assumption 2.1. Moreover, it is remarkable that if Σ_T is nonempty, then the set $h_0 \cdots h_{-T+1}(S)$, which contains the final values of the tracking processes started at $-T$, can have only finitely many values:

Lemma 3.1. *If Σ_T is nonempty, then $h_0 \cdots h_{-T+1}(S) \subset \Lambda_T$, where*

$$\Lambda_T := \{h_0 \cdots h_{-T+k+1}\, Z_{-T+k} \;:\; k = 1, \ldots, \sigma_T + 1\}. \tag{7}$$

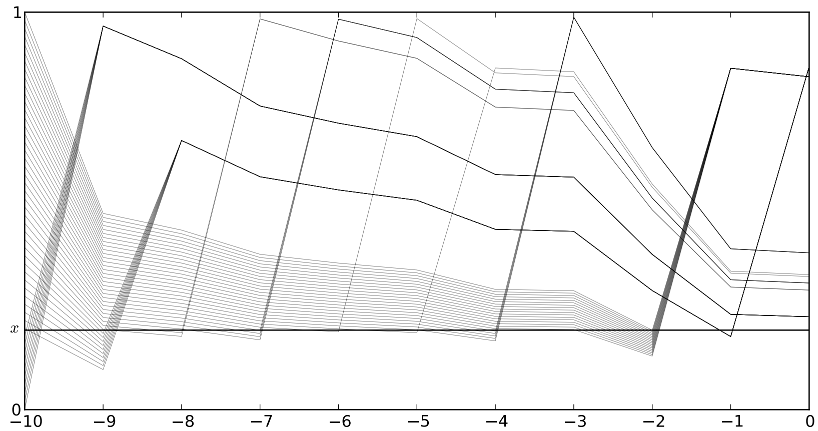

Figure 1: Tracking processes with $T = 10$ and $\sigma_T = 8$

The proof of lemma 3.1 is given in section 4. The intuition for the result can be obtained by considering figure 1. In the figure, $T = 10$. Tracking processes are plotted for 50 different initial values of $\phi \in S$. (Ideally, tracking processes would be plotted from every $\phi \in S$, but this is clearly impossible.) For this particular realization of shocks, the set Σ_T is nonempty because the process starting from 1 at time -10 falls below x at $t = -2$ (and hence $\sigma_T = 10 - 2 = 8$). As is clear from the figure, the fact that the process starting from 1 at time -10 falls below x at $t = -2$ implies that *all* tracking processes fall below x at least once between -10 and -2 (recall that the productivity of incumbents is monotone). Moreover, if any collection of tracking processes fall below x at some point in time t, they subsequently couple, taking the common value Z_{t+1} at $t+1$ and being equal from then on. As a result, by $t = -1$, there are at most $\sigma_T + 1 = 9$ distinct tracking processes. Their time zero values are included in the set Λ_T defined in lemma 3.1. In particular, Λ_T is the time zero values of the processes that start below x at dates $-10, -9, \ldots, -2$.

To see the importance of lemma 3.1, let $\{\Phi_t\}_{t\in\mathbb{Z}}$ be a stationary, doubly-indexed process on the same probability space as $\{(U_t, Z_t)\}_{t\in\mathbb{Z}}$ that obeys $\Phi_{t+1} = h_{t+1}\,\Phi_t$ for all $t \in \mathbb{Z}$. The common marginal distribution of Φ_t is μ. Since Φ_{-T} lies somewhere in S, we know that $\Phi_0 = h_0 \cdots h_{-T+1}\,\Phi_{-T} \in h_0 \cdots h_{-T+1}\,(S)$. Moreover, if the set Σ_T is nonempty, then lemma 3.1 yields the inclusion $h_0 \cdots h_{-T+1}\,(S) \subset \Lambda_T$, and Φ_0 lies in the finite observable set Λ_T. In particular, if Λ_T is a singleton, then the value of Φ_0 is revealed as the value of that singleton.

Figures 2 and 3 show simulations with successful and unsuccessful coalescence re-

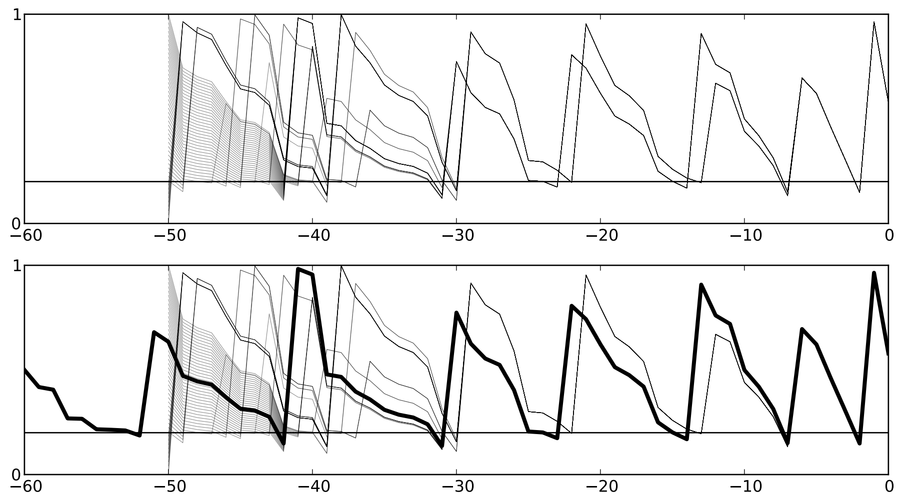

Figure 2: Successful coalescence from $T = 50$

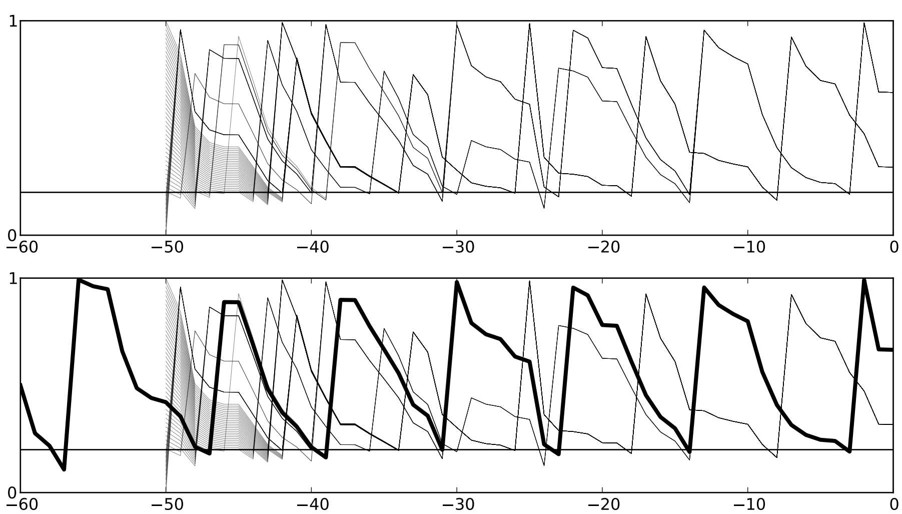

Figure 3: Failure of coalescence

spectively. In each figure, the top panel shows only the tracking processes. (As with figure 1, the full continuum of tracking processes cannot be plotted, so we show only 50.) The bottom panel shows the tracking processes and the path of $\{\Phi_t\}$. In reality, the path of $\{\Phi_t\}$ is not observed. However, in figure 2, there is only one final, coalesced value, and Φ_0 must take this value. Hence Φ_0 is observed. On the other hand, in figure 3, Φ_0 is equal to one of two final values, and we have no way of identifying which one it is.

Now let us consider how to use our results to sample from μ by generating observations of Φ_0. In order to avoid conditioning on coalescence by a certain point in time, we wish to reveal the value of Φ_0 for every random seed. This can be done by fixing the seed, which determines the values of the shock processes, and then taking T larger and larger until coalescence occurs. Algorithm 1 gives details. The algorithm terminates with an exact draw from μ. Replication with independent shocks will generate independent draws.

Algorithm 1: Generates an exact draw from μ

```
fix T to be an integer greater than 1;
draw (U_0, Z_0), ..., (U_{-T+1}, Z_{-T+1}) independently from their distributions;
repeat
    compute the set Σ_T ;
    if Σ_T is nonempty then
        compute the set Λ_T ;
        if Λ_T is a singleton then
            set Φ_0 to be the value of that singleton ;
            break ;
        end
    end
    draw (U_{-T}, Z_{-T}) and append to list (U_0, Z_0), ..., (U_{-T+1}, Z_{-T+1}) ;
    set T = T + 1 ;
end
return Φ_0 ;
```

At this stage we do not know that the algorithm will terminate with probability one. This issue is central to the correctness of the algorithm because, as discussed above, the way we avoid conditioning is by revealing the value of Φ_0 for every random seed. We now show that probability one termination in finite time holds under the following

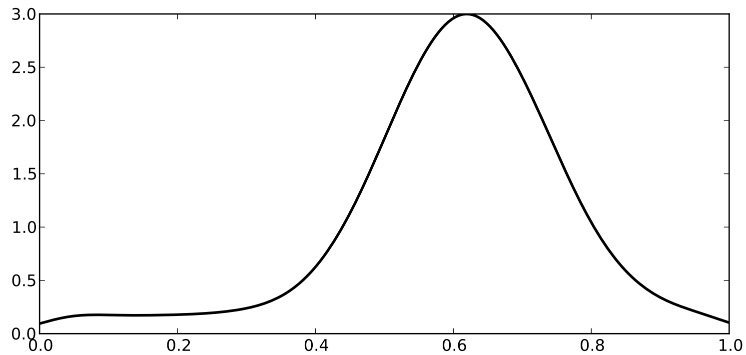

Figure 4: Stationary density

condition, which is satisfied by many standard distributions.

Assumption 3.1. If $G \subset S$ is a nonempty open set, then $\nu(G) > 0$.

Proposition 3.1. *Let* $T^* := \min\{T \in \mathbb{N} : \Sigma_T$ *is nonempty and* Λ_T *is a singleton* $\}$. *If assumption 3.1 holds, then there exists a* $\gamma \in (0, 1)$ *such that* $\mathbb{P}\{T^* > t\} \leq \gamma^t$. *In particular,* $\mathbb{P}\{T^* < \infty\} = 1$.

Note that proposition 3.1 not only gives probability one occurrence, but also provides the geometric rate $\mathbb{P}\{T^* > t\} = O(\gamma^t)$. The proof of proposition 3.1 is given in section 4.

The web page `https://github.com/jstac/hh_sampling` contains a simple implementation of algorithm 1. We tested the code by following Hopenhayn and Rogerson (1993) in taking the distribution ν for new entrants to be uniform, and the process for incumbents to be $\Phi_{t+1} = g_{t+1}\Phi_t = a + \rho\Phi_t + \epsilon_{t+1}$ where $\{\epsilon_t\}$ is IID with distribution $N(0, \sigma^2)$. To bound the process we added reflecting barriers at 0 and 1. The parameters were set to $a = 0.36$, $\rho = 0.4$ and $\sigma = 0.1$, while x was set to 0.49, so that approximately 40% of incumbents exit within 5 years (Hopenhayn, 1992, p. 1127). For these parameters, running the program on a standard workstation without parallelization produces about 36,000 independent draws from μ per second.[7]

Figure 4 shows the density computed from 36,000 observations combined with a standard nonparametric kernel density estimator (using a Gaussian kernel). Figure 5 shows a 95% confidence set for the cumulative distribution function corresponding to μ, based on the same observations and calculated using the Kolmogorov distribution of the sup

[7]Our workstation has a 2.67GHz Intel CPU and 4 gigabytes of RAM.

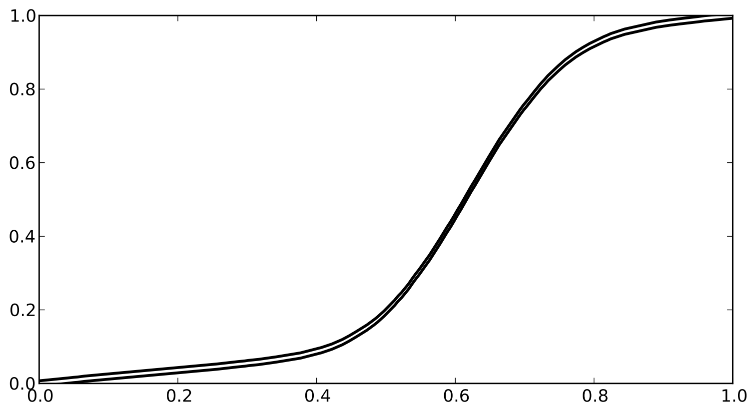

Figure 5: 95% confidence set for the stationary distribution

norm deviation between true and empirical cdfs. The Kolmogorov result is applicable here because the draws are exact and IID. The true distribution function lies entirely between the two bands with 95% probability.

4 Proofs

In the following proofs T is fixed, and we write σ for σ_T to simplify notation.

Proof of lemma 3.1. Let Σ_T be nonempty. As a first step, we show that if $\phi \geq x$, then there exists a $j \in \{1, \ldots, \sigma\}$ such that $h_{-T+j} \cdots h_{-T+1}\, \phi < x$. To see that this is so, fix $\phi \geq x$ and suppose that the statement fails. In other words, $h_{-T+j} \cdots h_{-T+1}\, \phi \geq x$ for $j \in \{1, \ldots, \sigma\}$. We know that if $y \geq x$, then $h_i\, y = g_i\, y$. It follows that $h_{-T+\sigma} \cdots h_{-T+1}\, \phi = g_{-T+\sigma} \cdots g_{-T+1}\, \phi$. But then

$$x \leq h_{-T+\sigma} \cdots h_{-T+1}\, \phi = g_{-T+\sigma} \cdots g_{-T+1}\, \phi \leq g_{-T+\sigma} \cdots g_{-T+1}\, 1 < x,$$

where the second inequality is due to monotonicity of g_i, and then third is by the definition of σ. Contradiction.

To complete the proof, pick any $\phi \in S$. Our claim is that $h_0 \cdots h_{-T+1}\, \phi \in \Lambda_T$. Suppose first that $\phi < x$. In this case we have $h_0 \cdots h_{-T+1}\, \phi = h_0 \cdots h_{-T+2}\, Z_{-T+1}$, which is an element of Λ_T. Next, suppose that $\phi \geq x$. In light of the preceding argument, there exists a $j \in \{0, \ldots, \sigma\}$ with $h_{-T+j} \cdots h_{-T+1}\, \phi < x$, and hence

$$h_{-T+j+1} \cdot h_{-T+j} \cdots h_{-T+1}\, \phi = Z_{-T+j+1},$$

from which we obtain

$$\begin{aligned} h_0 \cdots h_{-T+1} \phi &= h_0 \cdots h_{-T+j+2} \cdot h_{-T+j+1} \cdot h_{-T+j} \cdots h_{-T+1} \phi \\ &= h_0 \cdots h_{-T+j+2} Z_{-T+j+1}. \end{aligned}$$

Since $j \in \{0, \ldots, \sigma\}$, the right-hand side is an element of Λ_T. This completes the proof. □

Proof of proposition 3.1. Let n be an integer such that $F^n(x \mid 1) > 0$, existence of which is due to assumption 2.1. Fixing $j \in \mathbb{N}$, let

$$E_j := \{g_{-(j-1)n-1} \cdots g_{-jn} 1 < x\} \cap \{Z_{-(j-1)n-1} < x, \ldots, Z_{-jn} < x\}.$$

The events $\{g_{-(j-1)n-1} \cdots g_{-jn} 1 < x\}$ and $\{Z_{-(j-1)n-1} < x, \ldots, Z_{-jn} < x\}$ are independent because the first event depends only on $U_{-(j-1)n-1}, \ldots, U_{-jn}$ and the second depends only on $Z_{-(j-1)n-1}, \ldots, Z_{-jn}$. As a result,

$$\delta := \mathbb{P}(E_j) = F^n(x \mid 1) \nu([0, x))^n.$$

The constant δ is strictly positive as a result of assumption 3.1. We claim that if the event E_j occurs, then Σ_{jn+1} is nonempty and Λ_{jn+1} is a singleton. To simplify notation, we treat only the case of $j = 1$.

So suppose that E_1 occurs. Clearly Σ_{n+1} contains n, and hence is nonempty. To see that Λ_{n+1} is a singleton, observe that since $\sigma = \sigma_{n+1}$ is the smallest element of Σ_{n+1}, we must have $\sigma \leq n$. As a consequence,

$$\begin{aligned} \Lambda_{n+1} &= \{h_0 \cdots h_{-n+k} Z_{-n-1+k} \; : \; k = 1, \ldots, \sigma + 1\} \\ &\subset \{h_0 \cdots h_{-n+k} Z_{-n-1+k} \; : \; k = 1, \ldots, n + 1\}. \end{aligned}$$

We claim that on the set E_1 we have

$$h_0 \cdots h_{-n+k} Z_{-n+k-1} = Z_0 \quad \text{for any} \quad k \in \{1, \ldots, n+1\}. \tag{8}$$

To prove that (8) holds, observe that on E_1 the values $Z_{-1}, \ldots, Z_{-n}$ are all less than x. As a result, we have

$$\begin{aligned} h_0 Z_{-1} &= Z_0 \\ h_0 h_{-1} Z_{-2} &= h_0 Z_{-1} = Z_0 \\ h_0 h_{-1} h_{-2} Z_{-3} &= h_0 h_{-1} Z_{-2} = h_0 Z_{-1} = Z_0 \end{aligned}$$

and so on. Together, these equalities give (8). As a consequence, we conclude that Σ_{n+1} is nonempty and Λ_{n+1} is a singleton whenever E_1 occurs, and, more generally, Σ_{nj+1} is nonempty and Λ_{nj+1} is a singleton whenever E_j occurs. The events $E_1, E_2, \ldots$ are independent and have positive probability δ. Using the definition of T^*, we then have

$$\mathbb{P}\{T^* > nj\} = \mathbb{P}\{T^* \geq nj + 1\} \leq \mathbb{P} \cup_{i=1}^{j} E_i^c = (1-\delta)^j$$

for all $k \in \mathbb{N}$. Setting $\gamma := (1-\delta)^{1/n}$ gives the result stated in the proposition. □

References

Athreya, Krishna B and Örjan Stenflo (2003) "Perfect sampling for Doeblin chains," *Sankhyā: The Indian Journal of Statistics*, pp. 763–777.

Cooley, Thomas F and Vincenzo Quadrini (2001) "Financial markets and firm dynamics," *American Economic Review*, pp. 1286–1310.

Corcoran, JN and RL Tweedie (2001) "Perfect sampling of ergodic Harris chains," *The Annals of Applied Probability*, Vol. 11, pp. 438–451.

Cotroneo, Domenico, Roberto Natella, Roberto Pietrantuono, and Stefano Russo (2014) "A survey of software aging and rejuvenation studies," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, Vol. 10, p. 8.

Foss, Sergey G and RL Tweedie (1998) "Perfect simulation and backward coupling," *Stochastic models*, Vol. 14, pp. 187–203.

Hopenhayn, Hugo A (1992) "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica*, pp. 1127–1150.

Hopenhayn, Hugo and Richard Rogerson (1993) "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of political Economy*, pp. 915–938.

Huber, Mark (2003) "A bounding chain for Swendsen-Wang," *Random Structures & Algorithms*, Vol. 22, pp. 43–59.

Kijima, Shuji and Tomomi Matsui (2006) "Polynomial time perfect sampling algorithm for two-rowed contingency tables," *Random Structures & Algorithms*, Vol. 29, pp. 243–256.

Levin, David A, Yuval Peres, and Elizabeth L Wilmer (2009) "Markov chains and mixing times. With a chapter by James G. Propp and David B. Wilson," *American Mathematical Society, Providence, RI*.

Melitz, Marc J (2003) "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, Vol. 71, pp. 1695–1725.

Murdoch, Duncan J and Peter J Green (1998) "Exact sampling from a continuous state space," *Scandinavian Journal of Statistics*, Vol. 25, pp. 483–502.

Nishimura, Kazuo and John Stachurski (2010) "Perfect simulation of stationary equilibria," *Journal of Economic Dynamics and Control*, Vol. 34, pp. 577–584.

Propp, James Gary and David Bruce Wilson (1996) "Exact sampling with coupled Markov chains and applications to statistical mechanics," *Random structures and Algorithms*, Vol. 9, pp. 223–252.

Rust, John (1987) "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, pp. 999–1033.

Schumpeter, Joseph (1942) "Creative destruction," *Capitalism, socialism and democracy*.