DERIVATIVE ESTIMATES FROM SIMULATION OF CONTINUOUS-TIME MARKOV CHAINS

PAUL GLASSERMAN

Columbia University, New York, New York

(Received January 1989; revisions received April 1989, April, October 1990; accepted January 1991)

Countable-state, continuous-time Markov chains are often analyzed through simulation when simple analytical expressions are unavailable. Simulation is typically used to estimate costs or performance measures associated with the chain and also characteristics like state probabilities and mean passage times. Here we consider the problem of estimating *derivatives* of these types of quantities with respect to a parameter of the process. In particular, we consider the case where some or all transition rates depend on a parameter. We derive derivative estimates of the *infinitesimal perturbation analysis* type for Markov chains satisfying a simple condition, and argue that the condition has significant scope. The unbiasedness of these estimates may be surprising—a "naive" estimator would fail in our setting. What makes our estimates work is a special construction of specially structured parameteric families of Markov chains. In addition to proving unbiasedness, we consider a variance reduction technique and make comparisions with derivative estimates based on likelihood ratios.

Many types of systems studied through simulation can be modeled as continuous-time, countable-state Markov chains. Notable among these are examples arising in queueing and reliability. Frequently, such models become Markovian only when the natural state space is augmented to include (discrete) supplementary variables, as in the "method of stages." In such cases, the Markov property is preserved at the cost of a significant increase in the size and complexity of the state space. The resulting process may be intractable through ordinary numerical methods and simulation may become a computationally competitive alternative.

Typically, one is interested in the characteristics of not just one model but a class of models which may represent alternative designs under consideration or may reflect uncertainty about which model best fits an existing system. This paper considers a context in which a class of alternative Markov chains is described by a family $\{Q(\theta), \theta \in \Theta)\}$ of generator matrices that depend on a real or vector parameter θ. For example, θ could be a vector of service and arrival rates in a Markovian queueing network. The goal is to estimate derivatives with respect to θ of output characteristics associated with each $Q(\theta)$. Such derivatives can be used to drive stochastic approximation algorithms that optimize over θ. Another application is sensitivity analysis.

The most obvious way to estimate derivatives is to simulate at θ and $\theta + \Delta\theta$ and compare the results when $\Delta\theta$ is small; but such finite difference approximations are often statistically and numerically poorly behaved. Here, we are primarily concerned with estimates of the *infinitesimal perturbation analysis* type. (See Ho 1987 and Suri 1989 for overviews and extensive references.) Briefly, this method considers the effect on each sample path of a small change $\Delta\theta$ in the parameter. Consideration of these effects as $\Delta\theta \to 0$ leads to the use of almost-sure derivatives to estimate derivatives of expectations.

It is well known that perturbation analysis estimates work in some contexts and not others. It has been pointed out that there may be difficulties in applying perturbation analysis to "arbitrary" Markov chains if individual transition rates $q(x, y)$ change with θ. (See the discussions in, for example, Glasserman 1988a, Section 4, Heidelberger et al. 1988, Section 1, and Ho and Li 1988, Section 1). In fact, some restriction on the class of chains considered appears to be necessary. In this paper, we identify a simple sufficient condition which, if satisfied by a Markov chain, permits the use of perturbation analysis. The condition is so simple we paraphrase it here: Any two states with a common immediate predecessor have a common immediate successor.

At the heart of our analysis is a rather unusual construction of a parameteric family of Markov chains. Perturbing a transition rate $q(x, y)$ can have

Subject classifications: Simulation, statistical analysis of derivation estimates. Simulation, Markov processes: sensitivity analysis for.
Area of review: SIMULATION.

0030-364X/92/4002-0292 $01.25
© 1992 Operations Research Society of America

the effect of changing the state the process goes to upon leaving state x. Consequently, it would appear that even arbitrarily small changes in $q(x, y)$ could have drastic effects on the sample paths of the process. But the effect of parameter changes on individual sample paths depends entirely on how one constructs a parametric family of processes. The main contribution of this paper is to identify a construction (available for chains satisfying the condition above) in which changes in state transitions due to parameter changes are quickly corrected.

Using this construction, we obtain unbiased derivative estimates for a class of continuously accumulated, finite-horizon performance measures. However, our method cannot handle arbitrary random horizons; in particular, we must exclude the passage times that typically define a regenerative cycle. Our results do not, therefore, extend immediately to steady-state derivative estimation. While here we only consider continuous-time chains, our results are applicable in discrete time as well, because discrete-time chains can be embedded in continuous-time processes. Glasserman (1990) works out the details.

In Section 1, we describe our problem more precisely. Our special construction is outlined in Section 2, but the intricate details are left to Section 8. Based on this construction, we derive the algorithm presented in Section 3. Implementation of the algorithm does not require the special construction, so a reader uninterested in the detailed derivation may read Section 3 without reading Section 8. Section 4 considers the scope of our main condition and includes examples. Section 5 applies a variance reduction technique and shows how to modify the algorithm of Section 3 accordingly. In Section 6 we consider a different class of derivative estimates based on the *likelihood ratio method.* This method appears to be applicable with few restrictions. But our experience—consistent with that of others—is that when both methods are available, perturbation analysis yields better estimates. Some comparisons are made in Section 7. Longer proofs are collected at the end.

1. PRELIMINARIES

Following standard notation, we use Q to denote the generator matrix of a continuous-time, countable-state Markov chain $\{X_t, t \geq 0\}$. As usual, $q(x, y)$ is the instantaneous rate of transition of X from x to y. Our presentation is simplified if we assume from the outset that Q is conservative and has no absorbing or instantaneous states—i.e., for every state x:

$$0 < q(x) \equiv -q(x, x) = \sum_{y \neq x} q(x, y) < \infty.$$

We also assume that every row of Q has finitely many nonzero entries.

We consider a parametric family $\{Q(\theta), \theta \in \Theta\}$ of Markov chains in which the elements of Q depend continuously and, in fact, differentiably on θ. The argument θ will usually be suppressed. Henceforth, we take θ to be scalar. For vector parameters, gradients can be obtained by considering separately derivatives with respect to individual components. We fix the initial state, X_0, but any initial distribution independent of θ would do.

When we consider changes in θ, we understand that Q preserves its character as a conservative generator. Thus, when we increase $q(x, y)$ by δ, we get a corresponding decrease δ in $q(x, x)$. We also require that the (open) interval Θ be constrained so that, for all x, y, if $q(x, y)$ is positive at some $\theta \in \Theta$ it is positive throughout Θ. In other words, transition "arcs" can be neither created nor eliminated through changes in the parameter. This makes the family of measures on the path space induced by $\{Q(\theta), \theta \in \Theta\}$ mutually absolutely continuous.

The quantities whose derivatives we consider are variants of $J(\theta) = E_\theta[L]$ where

$$L = \int_0^T f(X_t)\, dt,$$

f is a bounded, real-valued function on the state space of X, and T is a stopping time. For most of the paper we restrict T to be either a fixed time or the epoch of the nth transition of X for some fixed n; in Section 4 we point out problems associated with the passage time to a state. Through choice of f and T, many quantities of interest can be obtained in this way. Of particular interest are the cases where f is the indicator of some set of states; where f is identically one (and T is stochastic); and where $f(x)$ represents some cost or reward associated with operating in state x. If Y_0, $Y_1, \ldots$ is the sequence of states visited by the process, we can also handle functions of the form $\sum_{i=0}^{n} f(Y_i)$, as explained in Glasserman (1990).

To estimate $dJ/d\theta = dE[L]/d\theta$ using the method of infinitesimal perturbation analysis, one must find a construction of processes $\{X_t(\theta), \theta \in \Theta\}$ (each $X_t(\theta)$ a Markov process with generator $Q(\theta)$) on a common probability space under which L is differentiable in θ with probability one. The resulting almost-sure derivative $dL/d\theta$ estimates the derivative of the expected value of L. This estimate is unbiased in case $E[dL/d\theta] = dE[L]/d\theta$. Different constructions lead to different almost-sure derivatives. It is possible for some to be biased and others not. Theorem 1 of Section 8.2 shows that the almost-sure derivative resulting from our construction is unbiased.

2. OUTLINE OF THE CONSTRUCTION

We construct a continuous-time Markov chain from a sequence of independent, unit mean exponential random variables $\{Z_i, i = 1, 2, \ldots\}$. There are many ways such a construction can be effected; before presenting a new way, we briefly describe a more obvious approach and why it fails.

Suppose that the process starts in state x_0. A standard construction assigns one Z_i to each pair of states (x_0, y_i) for which $q(x_0, y_i) > 0$ and simultaneously determines the first holding time and transition as follows. Think of assigning a "clock" to each possible transition (x_0, y_i); the clock runs at rate $q(x_0, y_i)$ and is initialized to Z_i. The ith clock is scheduled to run out at $Z_i/q(x_0, y_i)$. If it is the first to run out, then $Z_i/q(x_0, y_i)$ becomes the holding time in x_0, and the next state is y_i. After the transition, previously scheduled clocks are simply discarded, and new random variables from the sequence $\{Z_i\}$ are assigned to the possible transitions out of the new state. Clearly, under this construction, even a very small change in some $q(x_0, y)$ can completely change the sequence of states, making L discontinuous. Indeed, Heidelberger et al. show through explicit calculation that a derivative estimator based on this type of construction for a birth-death process converges to the wrong value.

The less obvious construction we propose modifies this approach, taking advantage of special structure. The condition we require is as follows:

Condition CM. *For any pair of states* y_1, y_2, *if there is a state* x *for which* $q(x, y_1) > 0$ *and* $q(x, y_2) > 0$, *then there must also be a state* x' *for which* $q(y_1, x') > 0$ *and* $q(y_2, x') > 0$.

If we think of the state space of a Markov chain as a directed graph with an arc from x to y whenever $q(x, y) > 0$, then the condition is that every pair of states with a common immediate predecessor must have a common immediate successor. In the case of a birth-death process, for example, the pairs of states with a common predecessor are all of the form $(x-1, x+1)$ and have a common successor x (which is also the common predecessor).

For a chain satisfying **CM**, if y_1 and y_2 have a common predecessor, let $K(y_1, y_2) = K(y_2, y_1)$ be a common successor. Generate the first holding time and state transition as described above; suppose these are determined by $Z_j/q(x_0, y_j)$. Following the transition to y_j, we *re-use* the residual time on the other clocks. The memoryless property applied to the other assigned Z_i's implies that, conditional on $Z_i/q(x_0, y_i) > Z_j/q(x_0, y_j)$, the quantity $Z_i/q(x_0, y_i) - Z_j/q(x_0, y_j)$ is exponentially distributed with its original mean, $q^{-1}(x_0, y_i)$. Thus, under the same condition, the residual time

$$Z_i - \frac{q(x_0, y_i)}{q(x_0, y_j)} Z_j$$

has a unit-mean exponential distribution. Assign this residual time on clock i to (y_j, x'), where $x' = K(y_i, y_j)$ (y_i and y_j have x_0 as a common predecessor). If it is the only clock assigned to this transition, run it at rate $q(y_j, x')$. If m (residual) clocks are assigned to the same transition (y_i, x') run each at rate $q(y_j, x')/m$. (Recall that the minimum of m independent exponential random variables with rate λ/m is exponential with rate λ.) If, following this reassignment, there is a state z with $q(y_j, z) > 0$ to which no clock has been assigned, then to each such z assign a new clock running at rate $q(y_j, z)$ and initialized to a freshly drawn, unit-mean exponential random variable. Repeat the whole procedure at each transition.

For example, consider the fragment of a state space depicted in Figure 1. If the process is started in state x_0, clocks 1 and 2 may be assigned to transitions (x_0, y_1) and (x_0, y_2), respectively. If clock 1 runs out first, then the process moves to state y_1 and clock 2 is reassigned to (y_1, x'). A new clock (clock 3) is assigned to (y_1, z). If the first transition is, instead, to state y_2, then clock 1 is reassigned to (y_2, x').

To see what makes this construction interesting, consider the effect of small changes in Q. Suppose initially the first transition is determined by

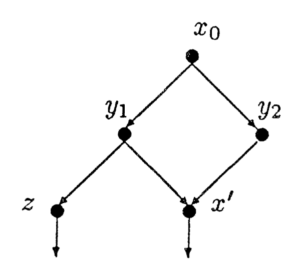

Figure 1. Fragment of a Markov chain.

$Z_j/q(x_0, y_j)$, and the next by

$$\left(Z_i - \frac{q(x_0, y_i)}{q(x_0, y_j)} Z_j\right) \bigg/ q(y_j, x')$$

where, as before, $x' = K(y_i, y_j)$. As $q(x_0, y_i)$ is increased, say by an amount $\delta > 0$, it may happen that $Z_j/q(x_0, y_j) > Z_i/(q(x_0, y_i) + \delta)$, in which case, the transition out of x_0 is to y_i rather than y_j. But, under our construction, after the transition to y_i

$$\left(Z_j - \frac{q x_0, y_j)}{q(x_0, y_i) + \delta} Z_i\right) \bigg/ q(y_i, x') \tag{1}$$

is assigned to $x' = K(y_j, y_i)$; and since initially $Z_j/q(x_0, y_j) < Z_i/q(x_0, y_i)$, (1) is close to zero whenever δ is small. For sufficiently small δ, the transition out of y_i is *necessarily* to x'. Thus, if initially the first three states were x_0, y_j, x', then for δ just large enough to introduce an order change, the new sequence will be x_0, y_i, x'; and, at that value of δ at which the sequence changes, the holding time in the second state is zero. (The reader can verify that neither of these conclusions holds under the first construction described above.) If, under the perturbation δ, the process reaches x' through y_i rather than y_j, we may still assign clocks to transitions out of x' in the same way we would have if the second state had been y_j. This ensures that (with high probability) the future evolution of the process is unaffected by the change in the second state. These arguments are made precise in Lemma 1.

3. THE ALGORITHM

By formalizing the scheme outlined above, we obtain a construction of Markov chains with generator matrices $Q(\theta)$, for a neighborhood of θ values, all from the same sequence $\{Z_i\}$ of exponential random variables. From that construction, we obtain expressions for derivatives (with respect to θ) of quantities associated with the chains. The construction, and the associated derivatives, are rather complicated and involve some detailed bookkeeping for the assignment and reassignment of clocks. The resulting derivative estimation algorithm is, in contrast, relatively simple; therefore, in this section we present only the algorithm. The construction and derivation are postponed to Section 8.

We need some notation. For each value of θ, associated with $Q(\theta)$ we have:

$X_t(\theta)$ = the state of the chain at time t;
$\tau_n(\theta)$ = the epoch of nth transition, $\tau_0(\theta) \equiv 0$;
$Y_n(\theta)$ = the nth state = $X_{\tau_n(\theta)}(\theta)$;
$N_t(\theta) = \sup\{n \geq 0: \tau_n(\theta) \leq t\}$;
= the number of transitions in $(0, t]$.

Our construction makes each $\{X_t(\theta), t \geq 0\}$ right-continuous, so $Y_n(\theta)$ is the state just after the nth transition. When the parameter value is fixed or irrelevant, we often omit the argument θ.

Let f be a bounded, real-valued function defined on the state space of X. For real $T > 0$ and integer $n > 0$ define

$$L_1 = L_1(T) = \int_0^T f(X_t)\,dt$$

$$= \sum_{i=0}^{N_T-1} f(Y_i)[\tau_{i+1} - \tau_i] + f(Y_{N_T})[T - \tau_{N_T}]. \tag{2}$$

$$L_2 = L_2(n) = \int_0^{\tau_n} f(X_t)\,dt = \sum_{i=0}^{n-1} f(Y_i)[\tau_{i+1} - \tau_i]. \tag{3}$$

If X depends on θ, then so do L_1 and L_2. Suppose that L_1 and L_2 are differentiable at some θ. The Y_i's take values in a discrete set, so $df(Y_i(\theta))/d\theta$ can only be zero, wherever it exists. The same is true of N_T. Thus, by differentiating the right-most expressions in (2) and (3) and rearranging terms, we obtain

$$\frac{dL_1}{d\theta} = \sum_{i=1}^{N_T} \frac{d\tau_i}{d\theta}[f(Y_{i-1}) - f(Y_i)] \tag{4}$$

and

$$\frac{dL_2}{d\theta} = \sum_{i=0}^{n-1} f(Y_i)\left[\frac{d\tau_{i+1}}{d\theta} - \frac{d\tau_i}{d\theta}\right]. \tag{5}$$

From these expressions it is clear that the key to evaluating the derivatives of L_1 and L_2 is evaluating the derivatives $\{d\tau_i/d\theta, i = 1, 2, \ldots\}$ of the transition epochs.

Our algorithm computes, for each state x, a sequence $\{\delta_i(x), i = 1, 2, \ldots\}$ with the following interpretation: If $q(Y_{i-1}, x) > 0$, then at τ_{i-1} there is a clock associated with a transition from Y_{i-1} to x. This clock determines the time remaining until the next *scheduled* transition to x. (A scheduled transition may be cancelled by the occurrence of another transition.) Think of $\delta_i(x)$ as representing the derivative, at τ_{i-1}, of the scheduled transition to x. In particular, $\delta_i(Y_i)$ represents the derivative of the time of the next actual transition. By substituting $\delta_i(Y_i)$ for $d\tau_i/d\theta$ in (4) and (5), the algorithm below computes estimates $\hat{L}_1'$

and $\hat{L}_2'$. (See Sections 8.2 and 8.3 for a more precise discussion.)

Given $Y_0, Y_1, \ldots, Y_i$, define

$$S_i(x) = \{y: K(Y_i, y) = x, q(Y_{i-1}, y) > 0\};$$

then $S_i(x)$ is the set of states y such that a clock assigned to (Y_{i-1}, y) at τ_{i-1} would be reassigned to (Y_i, x) at τ_i, under our construction. The set $S_i(x)$ could be empty and could consist of just one element.

Algorithm 1

Step 1. Initialize by setting every $\delta_0(x) = 0$, $\hat{L}_1' = 0$, and $\hat{L}_2' = 0$.

Step 2. At the ith transition, update by setting

$$\delta_i(Y_i) = \delta_{i-1}(Y_i) - \frac{\tau_i - \tau_{i-1}}{q(Y_{i-1}, Y_i)} q'(Y_{i-1}, Y_i) \tag{6}$$

for every x with $q(Y_i, x) > 0$ and $|S_i(x)| > 0$

$$\delta_i(x) = \delta_i(Y_i) + \sum_{y \in S_i(x)} \frac{q(Y_{i-1}, y)}{q(Y_i, x)} \cdot \left(\delta_{i-1}(y) - \frac{\tau_i - \tau_{i-1}}{q(Y_{i-1}, y)} q'(Y_{i-1}, y) - \delta_i(Y_i)\right); \tag{7}$$

while if $q(Y_i, x) > 0$ and $|S_i(x)| = 0$, then

$$\delta_i(x) = \delta_i(Y_i). \tag{8}$$

Set $\hat{L}_1' = \hat{L}_1' + \delta_i(Y_i)[f(Y_{i-1}) - f(Y_i)]$ and $\hat{L}_2' = \hat{L}_2' + f(Y_{i-1})[\delta_i(Y_i) - \delta_{i-1}(Y_{i-1})]$.

Step 3. After N_T transitions, stop incrementing $\hat{L}_1'$. After n transitions, stop incrementing $\hat{L}_2'$.

The three cases (6)–(8) implicitly reflect how small "perturbations" propagate from one clock to another. Some intuition for the form of these rules is given in Section 8.2, following a more detailed description of how clocks are assigned and reassigned.

In Section 8.3 we show that under mild additional conditions on $\{Q(\theta), \theta \in \Theta\}$, Algorithm 1 produces unbiased estimators; i.e., $E[\hat{L}_i'] = dE[L_i]/d\theta$, $i = 1, 2$. See Corollary 1. It should be stressed that while the derivation and justification of this algorithm use our special construction, the *implementation* does not. The algorithm itself makes no reference to clocks and could be used with any simulation of $\{X_t, t \geqslant 0\}$.

The notation of the algorithm assumes a separate sequence $\{\delta_i(x)\}$ for each state x. But because the algorithm has only a one-step memory, it is never necessary to store more δ_i values than the number of possible next transitions, which is generally small. Even so, associating accumulators with states presupposes some enumeration of the state space, which can be burdensome. For physically meaningful systems—the kind usually simulated—simpler implementations may be possible. We return to this point in the next section when we consider Jackson networks.

Section 5 shows that the algorithm lends itself well to *discrete-time conversion*, yielding lower-variance estimators with only minor modification.

4. SCOPE OF CONDITION CM

Our algorithm only yields correct results—indeed, can only be implemented—on Markov chains satisfying condition **CM**; hence, we should investigate the scope of this condition. In this section, we discuss some examples that satisfy **CM** and make some general remarks on this condition.

Example 1. (Reversible Markov Chains). These chains are characterized by the existence of strictly positive numbers $v(x)$ satisfying $q(x, y)v(x) = q(y, x)v(y)$ whenever $q(x, y) > 0$. For such chains, any two immediate successors of x have x itself as a common immediate successor. Clearly, the condition of reversibility could be relaxed to $q(x, y) > 0 \Rightarrow q(y, x) > 0$.

Example 2. (Jackson Networks). Consider an open or closed Jackson network with a single class of customers. Let P be the routing matrix; P_{ij} is the probability that a customer leaving server i joins queue j. In the open case, add a fictitious node 0: P_{0i} is the probability that an arriving customer joins queue i, and P_{i0} is the probability that upon leaving server i, a customer leaves the network. Suppose that all $P_{ii} = 0$ (this represents no loss of generality when service is exponential). Taking the state, as usual, to be the vector of queue lengths, **CM** is satisfied if P satisfies an analog of **CM**: For any i, if there are j and k for which $P_{ij}P_{ik} > 0$, then there is an i' such that $P_{ji'}P_{ki'} > 0$. (Note that i, j, k or i' may be 0.) Letting e_i be the ith unit vector and e_0 the vector of all zeros, this condition on P guarantees that if $n_i > 0$ in state $\mathbf{n} = (n_1, \ldots, n_M)$, then $\mathbf{n} - e_i + e_j$ and $\mathbf{n} - e_i + e_k$ have $\mathbf{n} - e_i + e_{i'}$ as a common successor. The same condition on P applies when arrivals, service and routing are state-dependent, so long as the dependence on the state does not include shutting off an arrival stream or a busy server, or zeroing a formerly positive routing probability. However, **CM** is typically violated by networks with multiple customer classes. See, e.g., the example in Heidelberger et al.

The special structure of Jackson networks makes implementation of Algorithm 1 especially easy. The $\delta_i(x)$'s can be replaced with accumulators $\delta_i(j, k)$ for all pairs of nodes j, k with $P_{jk} > 0$. Every state transition involves exactly two nodes. (Arrivals and departures involve node 0.) If $x = Y_i - e_j + e_k$, then at the ith transition update $\delta_i(j, k)$ using the rule for $\delta_i(x)$ in Algorithm 1. A similar simplification is possible in certain reliability models and, more generally, anytime the state space is (roughly) a Cartesian product. Fox (1990) exploits the same structure in a different way.

Example 3. (Perturbing Routing Probabilities). Perhaps the most striking difference between the estimators derived here and earlier infinitesimal perturbation analysis algorithms is that our estimators can be used when θ parameterizes the routing matrix, and not just the service times. If $\mathbf{n} = (n_1, \ldots, n_M)$ with $n_i > 0$, and if μ_i is the service rate at i, then $q(\mathbf{n}, \mathbf{n} - e_i + e_j) = \mu_i P_{ij}$; every nonzero, off-diagonal element of the generator matrix is a product of a service rate and a routing probability. In parameterizing Q by θ, it makes no difference which factor (μ_i or P_{ij}) depends on θ. Of course, for the network to satisfy **CM**, P must satisfy an analogous common successor condition, as explained in Example 2; hence, this method is only applicable with specially structured networks.

Example 4. (Phase-Type Distributions). The utility of our algorithm is substantially enhanced by the use of phase-type distributions. For example, this allows certain queueing networks with nonexponential service and interarrival times. Represent a phase-type distribution by an initial distribution α and an absorbing generator R. (See Neuts 1981 for background.) A phase-type service time, lifetime, etc., is the time to absorption through R starting according to α. Without loss of generality, let R have a unique absorbing state 0.

CM imposes restrictions on R and α; the specific restrictions depend on the system. As a simple example, consider queues in series. Consider distributions (α, R) for which: a) there is just one r such that $\alpha_r > 0$; b) if $R_{s0} > 0$, then for all s', $s \neq s' \neq 0$, $R_{ss'} = 0$; and, c) R satisfies **CM**. If the interarrival times and all service times are drawn from this class of distributions, then **CM** is satisfied.

When **CM** is not satisfied, it is tempting to try to extend the applicability of our method by replacing the true generator Q with a modified generator Q^ϵ, in which transitions of rate ϵ have been added to satisfy the condition. If ϵ is small, one might hope that this would produce reasonable estimates because the process is not greatly changed. Though this idea may work in some cases, it may be problematic. Introducing new transitions may allow meaningless state sequences in a physically meaningful system. More fundamentally, even if we make the reasonable assumption that the corresponding performance measures $E[L^\epsilon]$ converge to $E[L]$ as $\epsilon \to 0$, it seems less reasonable to assume that $dE[L^\epsilon]/d\theta \to dE[L]/d\theta$. (The convergence of a sequence of functions does not imply convergence of their derivatives.) Also, since transition rates appear in denominators in Algorithm 1, the estimates may diverge when $\epsilon \to 0$. For the same reason, the algorithm may perform poorly when some transition rates are very small.

Different conditions corresponding to a different class of perturbation analysis derivative estimates can be obtained using the *generalized semi-Markov process* (GSMP) framework, which does not rely on exponential distributions. This is done in Glasserman (1988b, 1991). It is important to stress that the key condition in these references, though superficially similar to **CM**, does not coincide with **CM** even when restricted to Markovian GSMPs. Neither condition implies the other, though there is overlap. Even when both conditions are in force, the corresponding derivative estimates may be different. A detailed comparison is not possible without the extensive GSMP notation. For Jackson networks, the difference may be summarized as follows. Results in Glasserman (1988b, 1991) place essentially no restrictions on the routing topology (compare Example 2) at the expense of strong conditions on the state-dependence allowed. Also, these results require that the routing not depend on the parameter of differentiation.

Finally, let us point out a limitation (alluded to earlier) associated with the performance measures we may consider. We cannot, in general, replace the time horizon T or τ_n with the passage time to a state. Speaking loosely, what **CM** guarantees is that if a change in θ causes X to jump to the "wrong" state, under a sufficiently small change, the next jump will be to the "right" state. This allows the jump epochs τ_n to be continuous in θ; see Lemma 1. But this correction is too coarse for passage times. If a parameter perturbation changes the state of the process for even a very short time, an entrance to a specified state may be created or eliminated, thus introducing a discontinuity in the passage time to that state. A consequence of this limitation is that our results are not immediately extensible to regenerative simulation.

5. CONDITIONING TO REDUCE VARIANCE

Simulations of continuous-time Markov chains are candidates for application of a variance reduction technique introduced by Hordijk, Iglehart and Schassberger (1976) and generalized by Fox and Glynn (1986, 1990). The idea—sometimes called "discrete-time conversion"—is to replace calculation of some estimate, r, from simulation of X with calculation of another estimate, $\hat{r}$, from simulation of the embedded chain Y. These estimates are related by $\hat{r} = E[r \mid Y]$. Simulating Y instead of X eliminates the need to generate exponential state holding times. Taking a conditional expectation reduces variance (see Fox and Glynn for details). Thus, $\hat{r}$ is preferable to r provided it is no harder to compute. In this section, we apply this method to the derivative estimates $\hat{L}_1'$ and $\hat{L}_2'$ produced by Algorithm 1.

From the right-most expression in (3), it is easy to see that

$$E[L_2 \mid Y] = \sum_{i=0}^{n-1} f(Y_i)E[\tau_{i+1} - \tau_i \mid Y] = \sum_{i=0}^{n-1} f(Y_i)/q(Y_i).$$

Thus, calculation of $E[L_2 \mid Y]$ simply requires replacing the holding time in each state by the mean holding time in that state. Conversion of $\hat{L}_2'$ is similar, as we show below. The conditional expectation for L_1, which is harder, is handled in Fox and Glynn (1990). We adapt one of their methods, based on uniformization, to convert $\hat{L}_1'$.

Inspection of the recursion that defines $\delta_i(x)$ (Algorithm 1) reveals that we may write

$$\delta_i(Y_i) = \sum_{j=0}^{i-1} \psi_{i-j}(Y_j, \ldots, Y_i)(\tau_{j+1} - \tau_j)$$

for some function $\psi_1, \psi_2, \ldots$. The value taken by each ψ_j is completely determined by Y and is otherwise independent of $\tau_1, \tau_2, \ldots$. Similarly, we find that the quantity, $\hat{L}_2'$ computed by our algorithm can be expressed as

$$\hat{L}_2' = \sum_{i=0}^{n-1} \psi(Y_i, \ldots, Y_n)(\tau_{i+1} - \tau_i), \tag{9}$$

for some (other) function ψ. Thus,

$$E[\hat{L}_2' \mid Y] = \sum_{i=0}^{n-1} \psi(Y_i, \ldots, Y_n)/q(Y_i).$$

In other words, the conditioned estimator $E[\hat{L}_2' \mid Y]$ is evaluated using the algorithm for $\hat{L}_2'$, replacing each holding time, $\tau_{i+1} - \tau_i$ with its conditional mean, $q^{-1}(Y_i)$. Thus, we may simulate only the embedded chain Y and apply essentially the same algorithm as before.

The case of L_1 is different because the number of transitions N_T in $(0, T]$ is not determined by Y. If (as we assume) there is a finite upper bound q^* on all transition rates, we may follow Fox and Glynn (1990) and uniformize the process at rate q^*. (Uniformization restricts us to simulation and is not applicable if the chain is merely "observed".) Denote by N^* the Poisson process of jumps of the uniformized chain, and by Y_i^* the ith state under uniformization. Let Y_i be the ith distinct state, and let M_i be the number of visits to Y_i the uniformized chain makes before proceeding to Y_{i+1}. Given Y_i, we generate the geometric variate M_i in $O(1)$ time. Given N_T^*, the ith holding time of the uniformized chain has mean $T/(N_T^* + 1)$; see Fox and Glynn (1990).

Using the ψ_j's above and (4), we can write $\hat{L}_1'$ as

$$\sum_{i=1}^{N_T} [f(Y_{i-1}) - f(Y_i)]\delta_i(Y_i) = \sum_{i=1}^{N_T} [f(Y_{i-1}) - f(Y_i)] \sum_{j=0}^{i-1} \psi_{i-j}(Y_j, \ldots, Y_i)(\tau_{j+1} - \tau_j).$$

Therefore:

$$\begin{aligned} E[\hat{L}_1' \mid Y^*, N_T^*] &= \sum_{i=1}^{N_T} [f(Y_{i-1}) - f(Y_i)] \cdot \sum_{j=0}^{i-1} \psi_{i-j}(Y_j, \ldots, Y_i)E[\tau_{j+1} - \tau_j \mid Y^*, N_T^*] \\ &= \sum_{i=1}^{N_T} [f(Y_{i-1}) - f(Y_i)] \cdot \sum_{j=0}^{i-1} \psi_{i-j}(Y_j, \ldots, Y_i) \frac{M_i T}{N_T^* + 1}. \end{aligned}$$

In words, the algorithm is modified by first generating N_T^*, generating M_{i-1} as well as Y_i in Step 2, and replacing the holding times with their expectations conditioned on Y^* *and* N_T^*. Thus, in (7), we replace $\tau_i - \tau_{i-1}$ by $M_{i-1}T/(N_T^* + 1)$.

While conditioning $\hat{L}_2'$ is virtually guaranteed to reduce variance with no additional work, there is a tradeoff in converting $\hat{L}_1'$. Uniformization typically involves slightly more work, so this must be compensated by the variance reduction, as discussed in Fox and Glynn (1990). Counterparts of more intricate results in Fox and Glynn (1990) further reduce variance with only a slight effect on work.

6. DERIVATIVES BASED ON LIKELIHOOD RATIOS

Various authors (Glynn 1986, Reiman and Weiss 1989, and Rubinstein 1989) have proposed estimating derivatives from simulation through likelihood ratios. The idea is to view θ as a parameter not of the sample paths of a process (as we have viewed it so far), but of a measure on the space of sample paths. Changes in θ lead to changes in the measure which lead naturally to likelihood ratios.

The domains of problems with which "perturbation analysis" and "likelihood ratio" derivative estimates can be used overlap but neither contains the other. It is interesting to compare their performance in cases where both can be applied. Markov chains satisfying condition **CM** are one such case. (While it is customary to think of these as alternative methods, in Glasserman (1990) we show that, for Markov chains, likelihood ratio estimates are perturbation analysis estimates obtained through yet another special construction. L'Ecuyer (1989) and Rubinstein (1989) also make connections between the methods.)

Let X, Q, Y and $\{\tau_n\}$ be as before and let $Y_0 \equiv y_0$ be fixed. Using basic properties of Markov chains

$$P(\tau_1-\tau_0\in dt_0, Y_1=y_1, \ldots, \tau_n-\tau_{n-1}\in dt_{n-1}, Y_n=y_n)$$
$$= e^{-q(y_0)t_0}\, dt_0 q(y_0, y_1) e^{-q(y_1)t_1}\, dt_1 q(y_1, y_2) \ldots e^{-q(y_{n-1})t_{n-1}}\, dt_{n-1} q(y_{n-1}, y_n). \quad (10)$$

We wish to consider small changes—in fact, derivatives—with respect to transition rates. We begin by differentiating with respect to a single $q(x, y)$, with $q(x, y) > 0$, then generalize to the parametric case. Collecting terms depending on $q(x, y)$, we may rewrite (10) as

$$q(x, y)^{N_{xy}} e^{-q(x)T_x} \cdot R \quad (11)$$

where N_{xy} is the number of transitions (x, y) among $(y_0, y_1), (y_1, y_2), \ldots, (y_{n-1}, y_n)$, T_x is the sum of those t_i for which $y_i = x$, and R does not depend on $q(x, y)$. (The same expression holds if Y_0 is chosen from an initial distribution that does not depend on $q(x, y)$.) The "likelihood ratio" of $(y_1, t_1, \ldots, y_{n-1}, t_{n-1}, y_n)$ at $q(x, y) + \delta$ relative to $q(x, y)$ is, therefore:

$$\frac{(q(x,y)+\delta)^{N_{xy}} e^{-(q(x)+\delta)T_x} \cdot R}{q(x,y)^{N_{xy}} e^{-q(x)T_x} \cdot R} = \left(1+\frac{\delta}{q(x,y)}\right)^{N_{xy}} e^{-\delta T_x}.$$

Differentiating with respect to δ, we get

$$\frac{N_{xy}}{q(x, y)}\left(1 + \frac{\delta}{q(x, y)}\right)^{N_{xy}-1} e^{-\delta T_x} + \left(1 + \frac{\delta}{q(x, y)}\right)^{N_{xy}} (-T_x) e^{-\delta T_x}.$$

Evaluating this at $\delta = 0$ we get

$$\frac{N_{xy}}{q(x, y)} - T_x. \quad (12)$$

Taking a likelihood ratio directly in (10) leads to a different expression; see (14) and also Section 3 of Glynn and Iglehart (1989). Collecting terms as in (12) does not change the work required, but does yield an analog of the Reiman and Weiss estimator for Poisson processes. Their analysis shows that expectation can be taken outside the derivative; more precisely, it yields the following proposition.

Proposition 1. *Let N_{xy} be the number of jumps from x to y on $(0, T)$ and T_x the time spent in state x during the same interval; then*

$$\frac{dE[L_1]}{dq(x, y)} = E\left[L_1\left(\frac{N_{xy}}{q(x, y)} - T_x\right)\right].$$

Proof. The proof of Theorem 1 in Reiman and Weiss applies here almost word for word. The "amiability" condition required there can be satisfied by taking the f in their notation equal to $T + N_T$, and taking their β to be any upper bound on the $|f|$ in our notation.

The case of L_2 is similar. One form of the likelihood ratio derivative estimate is given by

$$L_2\left(\frac{N_{xy}}{q(x, y)} - T_x\right), \quad (13)$$

where, now, N_{xy} is the number of jumps from x to y in $(0, \tau_n]$, and T_x is the time spent in x during that interval. This is easily derived from (16) in Glynn and Iglehart.

When the entire generator Q depends on some parameter θ, if the absolute continuity condition of Section 1 is in effect, then the chain rule yields the estimate

$$\frac{d\tilde{L}_1}{d\theta} = L_1 \sum_{(x,y)} \left(\frac{N_{xy}}{q(x, y)} - T_x\right) q'(x, y),$$

where the sum runs over all pairs of states with $q(x, y) > 0$. If only finitely many pairs have $q'(x, y) \neq 0$ and if where derivatives are bounded uniformly in θ, x, y, then unbiasedness follows from

Proposition 1. Estimates for L_2 can be found by summing over pairs of states in the same way.

Derivative estimates based on likelihood ratios can also be converted to discrete time. Consider the estimate

$$\frac{d\tilde{L}_2}{d\theta} = L_2 \sum_{(x,y)} \left(\frac{N_{xy}}{q(x,y)} - T_x \right) q'(x,y)$$

based on (13). Without L_2, the right side could be rewritten as

$$\sum_{i=0}^{n-1} \frac{q'(Y_i, Y_{i+1})}{q(Y_i, Y_{i+1})} - \sum_{i=0}^{n-1} q'(Y_i)(\tau_{i+1} - \tau_i), \quad (14)$$

using the fact that $\sum_y q'(x, y) = q'(x)$. Multiplying by L_2 introduces terms involving $(\tau_{i+1} - \tau_i)(\tau_{j+1} - \tau_j)$. Each factor is exponentially distributed; given Y, the two factors are independent if $j \neq i$. Using these observations, we get

$$E\left[\frac{d\tilde{L}_2}{d\theta} \,\middle|\, Y\right]$$
$$= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{q'(Y_i, Y_{i+1})}{q(Y_i, Y_{i+1})} f(Y_j)/q(Y_j)$$
$$- \sum_{i=0}^{n-1} \sum_{j \neq i} q'(Y_i) f(Y_j)/q(Y_i) q(Y_j)$$
$$- 2 \sum_{i=0}^{n-1} q'(Y_i) f(Y_i)/q^2(Y_i).$$

This equation also uses the fact that the second moment of an exponential random variable is twice the square of its mean. Collecting terms we get

$$\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left(\frac{q'(Y_i, Y_{i+1})}{q(Y_i, Y_{i+1})} - \frac{q'(Y_i)}{q(Y_i)} \right) f(Y_j)/q(Y_j)$$
$$- \sum_{i=0}^{n-1} q'(Y_i) f(Y_i)/q^2(Y_i),$$
$$= \bar{L}_2 \sum_{x,y} \left(\frac{N_{xy}}{q(x,y)} - \bar{T}_x \right) q'(x,y)$$
$$- \sum_{i=0}^{n-1} q'(Y_i) f(Y_i)/q^2(Y_i) \quad (15)$$

with $\bar{L}_2 = E[L_2 \mid Y]$ and $\bar{T}_x = E[T_x \mid Y]$. The first term in this last expression is what would be obtained if in the original estimate every holding time $\tau_{i+1} - \tau_i$ were replaced by its conditional mean $1/q(Y_i)$. Thus, in this case (in contrast to that of Section 5) simply replacing holding times by means it not legitimate; an additional term must be computed.

It is interesting to note—and not hard to check—that in (15), likelihood ratio differentiation and discrete-time conversion "commute": The likelihood ratio derivative estimate, based on the process Y_n, applied to $E[L_2 \mid Y]$ is the same as the conditional expectation evaluated above. Indeed, the two terms on the left side of (15) correspond to the two terms in (4.5) of Glynn for discrete-time chains. In our application, the first term reflects the dependence of the embedded transition probabilities on θ, the second the dependence of the mean holding times.

To convert $d\tilde{L}_1/d\theta$ to discrete time we use uniformization, as we did with L_1', again adapting the method of Fox and Glynn (1990). Let N^* and Y^* be as in Section 5, let τ_i^* be the ith transition epoch for $i \leq N_T^*$, and let $\tau_{N_T^*+1}^* = T$. Given N_T^*, the jumps of N^* are uniformly distributed on [0, T]. Using this fact, it can be shown that for $i, j \leq N_T^*$, $j \neq i$:

$$E[\tau_{i+1}^* - \tau_i^* \mid N_T^*, Y^*] = \frac{T}{N_T^* + 1},$$

$$E[(\tau_{i+1}^* - \tau_i^*)^2 \mid N_T^*, Y^*] = \frac{2T^2}{(N_T^* + 1)(N_T^* + 2)},$$

and

$$E[(\tau_{i+1} - \tau_i)(\tau_{j+1} - \tau_j) \mid N_T^*, Y^*] = \frac{T^2}{(N_T^* + 1)(N_T^* + 2)}.$$

An argument similar to that used for L_2 yields

$$E\left[\frac{d\tilde{L}_1}{d\theta} \,\middle|\, N_T^*, Y^*\right]$$
$$= \bar{L}_1 \left(\sum_{i=0}^{N_T^*-1} \frac{q'(Y_i^*, Y_{i+1}^*)}{q(Y_i^*, Y_{i+1}^*)} \right)$$
$$- 2 \sum_{i=0}^{N_T^*} \sum_{j=i}^{N_T^*} f(Y_i^*) q'(Y_i^*) \frac{T^2}{(N_T^* + 1)(N_T^* + 2)},$$

where $q'(y, y) \equiv 0$ for all y, and, from Fox and Glynn (1990)

$$\bar{L}_1 = E[L_1 \mid N_T^*, Y^*] = \frac{T}{N_T^* + 1} \sum_{i=0}^{N_T^*} f(Y_i^*).$$

Fox and Glynn (1990) discuss several improvements to $\bar{L}_1$, which should be applicable to $E[d\tilde{L}_1/d\theta \mid N_T^*, Y^*]$ as well.

7. COMPARISONS AND COMPUTATIONAL EXPERIENCE

A complete comparison of the relative efficiencies of derivative estimates based on the methods of Sections 3 and 6 is well beyond the scope of this paper. Nevertheless, some comparisons can be made in special cases and some conclusions drawn based on computational experience. (Further comparisons are made in L'Ecuyer 1989 and Rubinstein 1989.) The single greatest advantage of the likelihood ratio method, which cannot be quantified, is the fact that it appears to be applicable (to Markov chains) with only mild restrictions.

Implementation of the infinitesimal perturbation analysis estimate required from 30% to 50% more lines of code than the likelihood ratio method and typically took approximately twice as long to run. (Neither of these observations applies to simulations driven by future-event schedules. Most of the overhead in the perturbation analysis calculations goes to keeping track of the clocks (implicitly via S_i). When a future-even schedule is used, clock information is already available.) The larger the number of possible transitions from each state, the greater the work required just to generate transitions and the smaller the relative additional work to calculate derivatives by either method. For finite chains, checking condition **CM** and building the common successor matrix K requires a one-time effort of $O(k^3)$, where k is the number of states. For large k, this is prohibitive. For highly structured chains of the type arising in applications (such as those in Section 4) it is often possible to check **CM** and specify K from a verbal description of the process. (For Jackson networks, constructing a common successor matrix for the routing matrix P is sufficient; see Section 4.) In general, entries of K may be generated only as needed.

These apparent computational advantages are, however, quickly overwhelmed by the large variance typical of the likelihood ratio method. Consider (12): its variance is $\mathrm{Var}(E[(12) \mid T_x]) + E[\mathrm{Var}((12) \mid T_x)] = E[\mathrm{Var}((12) \mid T_x)]$ because $E[(12) \mid T_x] = 0$. At the root of the problem is the fact that when $q(x, y) > 0$,

$$\mathrm{Var}((12) \mid T_x) = E\left[\left(\frac{N_{xy}}{q(x, y)} - T_x\right)^2 \,\middle|\, T_x\right] = \frac{T_x}{q(x, y)}$$

(see also Reiman and Weiss, Section VIII), because given T_x, N_{xy} is Poisson distributed with parameter $q(x, y)T_x$. Since, typically, $E[T_x]$ increases in the length of the simulation, so would the variance of the differentiated likelihood ratio (12).

In examining the effect of run length on variance it makes sense to consider, for example, $L_2(n)/n$ since this usually has variance $O(n^{-1})$ (whereas the variance of $L_2(n)$ may increase without bound). If $\sigma_1^2(n)$ and $\sigma_2^2(n)$, respectively, denote the variances of the perturbation analysis and likelihood ratio derivative estimates of $L_2(n)/n$, then *typically*

$$\sigma_1^2(n) \sim O(n^{-1}) \quad \text{and} \quad \sigma_2^2(n) \sim O(n). \tag{16}$$

For example, we can expect the variance of $L_2(n)$ to be $O(n)$; that of the differentiated likelihood ratio is often also $O(n)$ as noted above. Multiplying the differentiated likelihood ratio by $L_2(n)/n$ we expect to obtain an estimator with variance $O(n)$, because, for large n, $L_2(n)/n$ is roughly a constant. The comparison in (16) can be verified exactly in special cases, and is also consistent with computational experience.

As an illustrative example for experimental comparison we chose a finite state birth-death process. Following standard practice, we compare the estimators on the basis of their relative efficiency: Let σ_1^2 be the variance of a perturbation analysis estimate and σ_2^2 that of a corresponding likelihood ratio method, and let t_1 and t_2 be the expected time to generate one estimate under each method. The perturbation analysis estimate is more *efficient* than the likelihood ratio estimate if

$$e_1 = (\sigma_1^2\, t_1)^{-1} > (\sigma_2^2\, t_2)^{-1} = e_2.$$

This gives a basis for trading off variance against computational effort, and also allows us to check the benefit of discrete-time conversion. (When samples are generated in constant time t_i, $\sigma_i\, t_i$ is the normalization in the ordinary central limit theorem for estimate i. When generation times are random, as in the present setting, this notion of efficiency is justified via more general central limit theorems in Glynn and Whitt 1989, and Fox and Glynn 1990.)

Tables I and II summarize our computational experience in the case of a five state birth-death process with $\lambda_{i-1} = \mu_i = 1$, $i = 1, \ldots, 5$ and $f \equiv 1$, for different values of n, with the without discrete-time conversion. (The simulations were written in FORTRAN and run on an AT&T PC 6300 PLUS.) The tables show estimates (based on a thousand replications) of $dE[\tau_n]/d\lambda_1$ using perturbation analysis and likelihood ratios, with estimated half-widths of 95% confidence intervals in parentheses. (In measuring relative efficiency it makes no difference whether we consider τ_n or τ_n/n.) The last column shows the (estimated)

Table I
Results for Birth-Death Process

n	$\hat{\tau}_n'$ PA	$\hat{\tau}_n'$ LR	e_1/e_2
5	−0.675 (0.037)	−0.681 (0.221)	20.3
10	−1.29 (0.06)	−0.970 (0.543)	42.7
20	−2.18 (0.09)	−1.75 (1.28)	105
30	−3.14 (0.11)	−2.33 (2.20)	206
40	−3.99 (0.13)	−3.42 (3.23)	320
50	−4.99 (0.14)	−2.85 (4.35)	486
75	−7.21 (0.17)	−1.33 (7.98)	1080
100	−9.61 (0.20)	−5.35 (12.28)	1890

Table II
Results for Birth-Death Process, Converted to Discrete Time

n	$\hat{\tau}_n'$ PA	$\hat{\tau}_n'$ LR	e_1/e_2
5	−0.674 (0.022)	−0.699 (0.130)	22.4
10	−1.29 (0.05)	−1.17 (0.333)	29.8
20	−2.21 (0.07)	−2.19 (0.859)	74.6
30	−3.16 (0.09)	−2.92 (1.51)	140
40	−3.99 (0.11)	−3.36 (2.28)	240
50	−5.00 (0.12)	−3.90 (3.21)	373
75	−7.25 (0.15)	−4.80 (5.90)	844
100	−9.62 (0.17)	−7.11 (9.21)	1515

relative efficiencies; for each n, the number in the last column is how many times more efficient than the LR estimate is the corresponding PA estimate. These suggest that, very quickly, the smaller variance associated with the perturbation analysis estimate dominates the increased computational requirement. (The average CPU times per thousand replications per n for the two methods are 3.6 and 1.9 seconds, and 3.2 and 1.7 seconds with discrete-time conversion.) Similar results were observed with other examples.

Our experience (not shown explicitly in the tables) is that the direct benefit in computation time from discrete-time conversion is modest—about 10%. The direct effect on variance (also modest) is indicated by the slightly tighter confidence intervals. (Theoretically, however, the benefit from conversion can be arbitrarily large or small; see Example 1 of Fox and Glynn 1990.) Overall, our experience is that converting to discrete-time benefits estimates of L_1 and L_2 more than estimates of their derivatives, and benefits the likelihood ratio method a bit more than perturbation analysis (compare the last columns of the two tables). However, it is unclear if such conclusions can be drawn more generally.

8. DERIVATION OF THE ESTIMATOR

We now detail the construction outlined in Section 3, and show how Algorithm 1 follows from this construction.

8.1. The Construction

We construct a Markov process with generator $Q(\theta)$ for a range of θ values. There are two parts: We first follow the steps outlined in Section 3 at a "nominal" value θ_0. This is the value at which the actual simulation would take place. We then "perturb" the evolution of the nominal process to obtain a Markov process with generator, say $Q(\theta_0 + h)$.

We begin at some θ_0 but suppress the parameter. As before, we denote by $\{X_t,\ t \geq 0\}$ the constructed Markov chain. The following summarizes the rest of our notation:

τ_n = the epoch of the nth transition;
Y_n = the nth state $= X_{\tau_n}$;
J_n = the index of clock that determines the nth transition;
$r_n(j)$ = the arc to which clock j is assigned at τ_n;
$c_n(j)$ = the time remaining on clock j at τ_n;
C_n = the set of clocks assigned at τ_n;
N_n = the number of clocks assigned or used up to τ_n;
$= \max\{j \in C_n \cup \{J_1, \ldots, J_n\}\}$;
$A(x)$ = the set of arcs (x, y) with $q(x, y) > 0$;
$d(a)$ = the destination of arc a—e.g., $d((x, y)) = y$.

Assume for the moment that for every x, if $K(x, y) = K(x, z)$, then $y = z$. Then the trajectories of X_t are determined as follows. Initialize by setting $\tau_0 = 0$; sampling Y_0 from an initial distribution; setting $N_0 = |A(Y_0)|$ and $C_0 = \{1, \ldots, N_0\}$. If $j \in C_0$, then $c_0(j) = Z_j$. Assign (using r_0) the clocks in C_0 to arcs in $A(Y_0)$ arbitrarily (but one to one). Then repeat the following recursion:

$$\tau_{n+1} = \tau_n + \min\{c_n(j)/q(r_n(j)): j \in C_n\}; \tag{17}$$

$$J_{n+1} = \min\{j \in C_n: c_n(j)/q(r_n(j)) = (\tau_{n+1} - \tau_n)\}. \tag{18}$$

(We allow the possibility that two clocks run out at the same time. Although such an event has probability zero for any given Q, the possibility must be admitted as Q is varied.)

$$Y_{n+1} = d(r_n(J_{n+1})). \tag{19}$$

Let

$$N_{n+1} = N_n + |A(Y_{n+1})| - |C_n| + 1, \tag{20}$$

and

$$C_{n+1} = C_n - \{J_{n+1}\} \cup \{N_n + 1, \ldots, N_{n+1}\}. \tag{21}$$

(The right-most set in (21) is empty if $N_{n+1} = N_n$.) For $j \in C_{n+1} \cap C_n$;

$$c_{n+1}(j) = c_n(j) - q(r_n(j))[\tau_{n+1} - \tau_n] \tag{22}$$

and

$$r_{n+1}(j) = (Y_{n+1}, K(Y_{n+1}, d(r_n(j)))); \tag{23}$$

while any $j \in C_{n+1} \backslash C_n$ is assigned a $c_{n+1}(j)$ arbitrarily from $\{Z_{N_{n+1}}, \ldots, Z_{N_{n+1}}\}$ and an $r_{n+1}(j)$ from the unassigned elements of $A(Y_{n+1})$.

If we allow $K(x, y) = K(x, z)$ even when $y \neq z$, then it may be necessary to assign multiple clocks to the same transition; upon entry to x, any clocks previously assigned to y and z would be reassigned to $K(x, y) = K(x, z)$. This case works the same way, except that when m clocks are assigned to an arc, (x, y), each is run down at rate $q(x, y)/m$. This makes it possible in (23) to have $d(r_n(j)) = Y_{n+1}$ (if j and J_{n+1} are assigned to the same arc). To cover this case, every $K(y, y)$ may be defined to be an arbitrary element of $A(y)$. In other words, any clocks other than J_{n+1} previously assigned to (Y_n, Y_{n+1}) may be assigned arbitrarily to some transition out of Y_{n+1}. (Notice also that, by construction, once two clocks are assigned to a common transition, they remain assigned to a common transition: in (23), $r_n(j) = r_n(j')$ implies $r_{n+1}(j) = r_{n+1}(j')$.)

Let $X_t = Y_n$ on $[\tau_n, \tau_{n+1})$. Using basic properties of the exponential distribution (mentioned in Section 2), one verifies the following proposition.

Proposition 2. *X_t is a continuous-time Markov chain with generator Q.*

We now describe how the "perturbed" $\theta_0 + h$ process is obtained from the "nominal" θ_0 process, denoting these by $X_t(\theta_0)$ and $X_t(\theta_0 + h)$. This step is needed to complete the construction, but does not directly enter into the derivation of Algorithm 1. The evolution of $X_t(\theta_0 + h)$ begins by following the construction above (but driven by $Q(\theta_0 + h)$ rather than $Q(\theta_0)$). As h is increased (or decreased) there may be a change in which a clock triggers a transition. Let $\tilde{N}$ be the number of transitions, either n or N_T. It may happen that, for some $i \leqslant \tilde{N}$, $J_i(\theta_0 + h) \neq J_i(\theta_0)$. This potentially introduces a change in the state sequence. Let h be just large enough for such an order change to occur, and let i be, in fact, the smallest index for which $J_i(\theta_0 + h)$ has a discontinuity at $\theta_0 + h$. At this point, we distinguish two cases.

Case 1. $Y_{i+1}(\theta_0 + h) = Y_{i+1}(\theta_0)$. This is the case if $J_i(\theta_0)$ and $J_{i+1}(\theta_0)$ have simply changed order; the construction has brought us back to the "right" state. There is, however, a possible discrepancy in the assignment of clocks. To correct it, simply set $C_{n+1}(\theta_0 + h) = C_{n+1}(\theta_0)$, set $N_{n+1}(\theta_0 + h) = N_{n+1}(\theta_0)$, and for any $j \in C_{n+1}(\theta_0)$, set $r_{n+1}(j, \theta_0 + h) = r_{n+1}(j, \theta_0)$. Now let the chain continue to evolve according to (17)–(23), driven by $Q(\theta_0 + h)$.

Case 2. $Y_{n+1}(\theta_0 + h) \neq Y_{n+1}(\theta_0)$. For this case to occur, more than two clocks must run out at the same time as J_n. As we will see, this case is actually negligible. In order, however, that $X_t(\theta_0 + h)$ be defined even in this case, we simply allow it to continue to evolve according to (17)–(23), but driven by $Q(\theta_0 + h)$.

The advantage of this construction is made precise in Lemmas 1 and 2, below. To state them, fix a real number $T > 0$ and an integer $n > 0$ and, as above, the $\tilde{N}$ denotes either n or $N_T(\theta_0)$. The first result is proved in Section 9.

Lemma 1. *Suppose that* **CM** *holds and the elements of $Q(\theta)$ are continuously differentiable functions of θ. Suppose there are positive constants B, q_* and q^* independent of θ, such that at every θ in Θ, for all x and y, $|q'(x, y)| \leqslant B$, and if $q(x, y) > 0$, then*

$$0 < q_* < q(x, y) < q(x) < q^* < \infty.$$

For any $\theta_0 \in \Theta$, under the construction above, the following hold with probability $1 - O(h^2)$: i) Every τ_i, $i \leqslant \tilde{N}$ is continuous in θ throughout $(\theta_0 - h, \theta_0 + h)$. ii) For every $i \leqslant \tilde{N}$, at any discontinuity of Y_i in $(\theta_0 - h, \theta_0 + h)$, $\tau_{i+1} = \tau_i$.

From Lemma 1 we get:

Lemma 2. *L_1 and L_2 are, with probability $1 - O(h^2)$, continuous functions of θ throughout $(\theta_0 - h, \theta_0 + h)$.*

Proof. Rewrite L_2 as

$$L_2 = \sum_{i=0}^{n-1} f(Y_i)[\tau_{i+1} - \tau_i]. \tag{24}$$

Consider the set of values of $\{Z_i\}$ for which the conclusion of Lemma 1 holds. On this set, part *i* of Lemma 1 implies that the only possible points of discontinuity in $(\theta_0 - h, \theta_0 + h)$ are those of some Y_i. Part *ii* guarantees that L_2 is continuous even at such points. Thus, the probability that L_2 is continuous throughout $(\theta_0 - h, \theta_0 + h)$ is as great as the probability that *i* and *ii* hold. For L_1, there is an additional case corresponding to a possible discontinuity in $Y_{N_T} = X_T$; see the right-most expression in (2). But a discontinuity in Y_{N_T} occurs only when there is transition right at T, in which case $[T - \tau_{N_T}] = 0$, so L_1 is still continuous.

8.2. The Estimator

We now show how to calculate sample path derivatives resulting from the construction of the previous section. The result, as proved in Theorem 1, is an unbiased derivative estimate. This form of the estimate requires keeping track of the evolution of the clocks. In Section 8.3, we modify the estimate to make it applicable without reference to clocks, and obtain Algorithm 1.

For notational convenience, we use $q_i(j)$ for the rate at which clock j is run down during $[\tau_i, \tau_{i+1})$; thus, $q_i(j)$ is $q(r_i(j))$ divided by the number of clocks in C_i assigned to $(Y_i, r_i(j))$. With each clock j associate an accumulator $D(j)$ which takes the value $D_i(j)$ at τ_i. Initialize every $D_0(j)$ to zero. At the ith transition update according to

$$D_i(J_i) = D_{i-1}(J_i) - \frac{\tau_i - \tau_{i-1}}{q_{i-1}(J_i)} q'_{i-1}(J_i); \tag{25}$$

if $j \in C_i \cap C_{i-1}$, then

$$D_i(j) = \left(\frac{q_{i-1}(j)}{q_i(j)}\right)\left(D_{i-1}(j) - \frac{\tau_i - \tau_{i-1}}{q_{i-1}(j)} q'_{i-1}(j)\right) + \left(1 - \frac{q_{i-1}(j)}{q_i(j)}\right) D_i(J_i); \tag{26}$$

and if $j \in C_i \backslash C_{i-1}$ then,

$$D_i(j) = D_i(J_i). \tag{27}$$

Roughly speaking, $D_i(j)$ is the infinitesimal delay in the time clock j is scheduled to run out. Suppose that a clock running in isolation at rate q takes Δ time units to run out. Then the derivative of the time it takes to run out this clock is $-(\Delta/q)q'$. Thus, the term added on the right in (25) is the (negative) delay in the time J_i runs out, due solely to an increase in $q_{i-1}(J_i)$. This is the "perturbation" in clock J_i generated by a change in $q_{i-1}(J_i)$. At the ith transition, clocks in $j \in C_i \cap C_{i-1}$ are reassigned; the factor $q_{i-1}(j)/q_i(j)$ in (26) rescales the perturbation in clock j to its new rate. A delay in the consumption of clock J_i delays the reassignment of clock j, causing it to run longer at rate $q_{i-1}(j)$ and shorter at rate $q_i(j)$. The last term in (26) captures this effect. Finally, if $j \in C_i \backslash C_{i-1}$, then j is set when J_i runs out, so any delay in J_i is directly propagated to j. This explains (27).

The precise result, proved in Section 9, is

Proposition 3. *With probability one, for every* $n \geqslant 0$, $d\tau_n/d\theta = D_n(J_n)$.

From the construction of the previous section it is clear that at any θ, every τ_n is differentiable with probability one; the only points where τ_n may not be differentiable are the points of discontinuity of J_n—where two clocks run out at the same time. Also, every Y_n is piecewise constant in θ. From the right-most expressions in (2) and (3) we conclude that, at each θ, L_1 and L_2 are differentiable with probability one, and their derivatives are indeed given by (4) and (5). Hence, these derivatives can be computed from the $\{D_n(j)\}$ as the process evolves.

We now come to our main result. As discussed in Section 1, we restrict attention to a set Θ throughout which every $\mathbf{1}\{q(x, y) > 0\}$ is a constant function of θ; i.e., no positive transition probabilities are created or eliminated through changes in the parameter. The following is proved in Section 9.

Theorem 1. *Under the conditions of Lemma* 1, *at any* $\theta \in \Theta$,

$$E\left[\frac{dL_i}{d\theta}\right] = \frac{dE[L_i]}{d\theta}, \quad i = 1, 2.$$

8.3. Simplification of the Estimator

One way to implement our derivative estimate generates sample paths of a Markov process using the construction in Section 8.1 and applies the recursion for $\{D_n(j)\}$ of Section 8.2. However, it is neither necessary nor particularly desirable to use the construction in applying the estimate. Indeed, if the algorithm is to be applied to observation of a real system, then using our special construction is not even an option.

Algorithm 1 keeps track only implicitly of how the clocks would be evolving if they were driving the process. It depends only on $X = \{X_t, t \geqslant 0\}$, and not on the mechanism that generates X. In fact, we now show that the $\hat{L}'_i$, $i = 1, 2$ computed by Algorithm 1 are just $E[dL_i/d\theta \mid X]$, $i = 1, 2$. Once conditioned on X, the estimators depend only on X, and can be applied to sample paths observed or simulated by any means. In particular, they can be coupled with specialized, efficient methods for simulating Markov chains (such as the one described in Fox 1990) which are otherwise incompatible with our construction.

If, for some Q, it is never necessary to assign multiple clocks to a single arc in the construction of Section 8.1 (e.g., if $K(x, y) = K(x, z) \Rightarrow y = z$), then the evolution of the clocks is, in fact, completely determined by X, and $E[dL_i/d\theta \mid X] = dL_i/d\theta$, $i = 1$, 2. In this case, it is possible to keep track of the $D_n(j)$ from observation of X. More generally, we get

$$E\left[\frac{dL_1}{d\theta} \,\middle|\, X\right] = \sum_{i=1}^{N_T} [f(Y_{i-1}) - f(Y_i)] E[D_i(J_i) \mid X],$$

and an analogous expression for L_2. But once a set of clocks is assigned to a common arc, they remain assigned to a common arc, and become indistinguishable. By construction, given X, J_i is equally likely to be any of the clocks assigned to (Y_{i-1}, Y_i) at τ_i; so to get $E[D_i(J_i) \mid X]$ we *average*.

Under the initialization of our construction, at time

0 there is just one clock assigned to (Y_0, Y_1); thus:

$$E[D_1(J_1) \mid X] = D_1(J_1) = -\frac{\tau_1 - \tau_0}{q(Y_0, Y_1)} q'(Y_0, Y_1)$$

the second equality following from (25). Since the $D_n(j)$'s enter linearly in the recursion of Section 8.2, we can apply the recursion to their conditional expectations to get $E[D_i(J_i) \mid X]$. This is precisely what Algorithm 1 does; $\delta_i(Y_i)$ is just $E[D_i(J_i) \mid X]$. The correspondence between (6) and (25), and between (8) and (27) is clear. To compare (7) and (26), recall that $S_i(x)$ is the set of states y, such that a clock assigned to (Y_{i-1}, y) at τ_{i-1} is reassigned to (Y_i, x) at τ_i. Thus, (7) averages (26) over clocks reassigned to the same transition. The $|S_i(x)|^{-1}$ that makes the sum more obviously an average is cancelled by the rate in the denominator, because if $r_i(j) = x$, then $q_i(j) = q(Y_i, x)/|S_i(x)|$.

Since $E[\hat{L}_i'] = E[E[dL_i/d\theta \mid X]] = E[dL_i/d\theta]$, $i = 1, 2$, from Theorem 1 we have a corollary.

Corollary 1. *Under the conditions of Lemma* 1, *Algorithm* 1 *produces unbiased estimators of* $dE[L_i]/d\theta$, $i = 1, 2$.

9. PROOFS

As in Section 8, here we use $q_i(j)$ for the rate at which clock j is run during $[\tau_i, \tau_{i+1})$.

Proof of Lemma 1. The proof proceeds in two parts. We first restrict attention to the case where for at most one $i < \tilde{N}$, J_i or J_{i+1} has a discontinuity in $(\theta_0 - h, O_0 + h)$, and show that i and ii hold in this case. We then verify that this case indeed has probability $1 - O(h^2)$. We refer to this as the *single order-change* case.

From the construction, it is clear that all τ_k and Y_k, $k \leq i$ are continuous at a point $\theta_0 + h$ if $J_1, \ldots, J_i$ are. As h is increased, let η be the smallest index such that J_η has a discontinuity in $(\theta_0 - h, \theta_0 + h)$, with $\eta = \infty$ if no such discontinuity arises. If $\eta > \tilde{N}$, then i and ii hold automatically. Thus, suppose $\eta \leq \tilde{N}$, and suppose, also, that the discontinuity in J_η occurs right at $\theta_0 + h$. In order that J_η be discontinuous, there must be two clocks j and k which run out at the same time; i.e.,

$$c_{\eta-1}(j)/q_{\eta-1}(j) = c_{\eta-1}(k)/q_{\eta-1}(k) = \min\{c_{\eta-1}(i)/q_{\eta-1}(i): i \in C_{\eta-1}\}, \quad (28)$$

with all quantities evaluated at $\theta_0 + h$. Comparing (28) and (17) we see that regardless of any discontinuity in J_η, at $\theta_0 + h$,

$$\tau_\eta = \tau_{\eta-1} + c_{\eta-1}(j)/q_{\eta-1}(j) = \tau_{\eta-1} + c_{\eta-1}(k)/q_{\eta-1}(k)$$

is continuous. Shortly, we show that this extends to every τ_i, $i \leq \tilde{N}$.

Next, note that a discontinuity in J_η may introduce one in Y_η—in particular, Y_η would jump from $d(r_{\eta-1}(j))$ to $d(r_{\eta-1}(k))$ if J_η jumped from j to k (see (19)). But regardless of this possibility, (28) and (22) together imply that there will be at least one clock in C_η with a residual time of zero. In light of (17), this means that $\tau_{\eta+1} = \tau_\eta$. Next, we extend this to every $i \leq \tilde{N}$.

In the single-order change case we are considering, at most the two clocks j and k run out at the same time (at $\tau_\eta = \tau_{\eta+1}$). Regardless of the order in which they occur, $Y_{\eta+1}$ is the same state—namely, $K(d(r_{\eta-1}(j)), d(r_{\eta-1}(k)))$. Hence, Case 1 of the construction of $X_t(\theta_0 + h)$ is in effect, so $N_{\eta+1}$, $C_{\eta+1}$ and every $c_{\eta+1}(j')$ with $j' \in C_{\eta+1}$ remain unchanged at a discontinuity of J_η and Y_η. Thus, every τ_i and Y_i, $i > \eta + 1$, are continuous at $\theta_0 + h$ if $J_{\eta+2}, \ldots, J_i$ are. But in the single-order change case, every J_i, $\eta + 1 < i \leq \tilde{N}$ is, in fact, continuous. Thus, i and ii hold.

We now verify that exceptions to the single-order change case occur with probability at most $O(h^2)$. There are two types of exceptions. It could happen that through an increase in h, more than two clocks run out at the same time; or, it could happen that there are two Y_i, $i \leq \tilde{N}$, with discontinuities in $(\theta_0 - h, \theta_0 + h)$. We refer to either of these cases as *multiple order changes*.

To consider the probabilities of these events, we first need to bound the change in the τ_i due to changes in h. Let $\tilde{h}$ be the infimum over h for which a multiple order change occurs in $(\theta_0 - h, \theta_0 + h)$ ($\tilde{h}$ is strictly positive with probability one). For $h < \tilde{h}$, the $\tau_i(\theta_0 + h)$ are continuous and have (at least) left and right-hand derivatives. Taking the larger of these at each point, the bound (38) proved below (which is valid throughout $(\theta_0 - \tilde{h}, \theta_0 + \tilde{h})$) yields

$$\left| \frac{d\tau_i}{d\theta} \right| \leq (B\rho^i/q_*)\tau_i, \quad (29)$$

where ρ is a positive constant. Integrating, we get, for $h < \tilde{h}$,

$$e^{-hB\rho^i/q^*}\tau_i(\theta_0) \leq \tau_i(\theta_0 + h) \leq e^{hB\rho^i/q^*}\tau_i(\theta_0);$$

so,

$$|\tau_i(\theta_0 + h) - \tau_i(\theta_0)| \leq (e^{hB\rho^i/q^*} - 1)\tau_i(\theta_0). \quad (30)$$

Using the mean value theorem (or expanding the exponential in a Taylor series), the right side can be replaced with $he^{hB\rho^i/q^*}\tau_i$. Since we are interested only in the limit as $h \to 0$, we may, without loss of generality, consider only $h < 1$. Thus, the bound becomes $he^{B\rho^i/q^*}\tau_i$, which we rewrite as $h\beta_i\tau_i$ by letting $\beta_i = e^{B\rho^i/q^*}$. For J_i and J_{i+1} to run out at the same time at $\theta_0 + h$, $h \leq \tilde{h}$, the epochs τ_i and τ_{i+1} must coincide. For this to happen, the net movement of τ_i and τ_{i+1} as the parameter varies from θ_0 to $\theta_0 + h$ must exceed their original separation, $\tau_{i+1}(\theta_0) - \tau_i(\theta_0)$. That is, we must have

$$\begin{aligned}|\tau_{i+1}(\theta_0) - \tau_i(\theta_0)| &\leq |\tau_{i+1}(\theta_0+h)-\tau_{i+1}(\theta_0)| + |\tau_i(\theta_0+h)-\tau_i(\theta_0)|\\ &\leq \beta_i h\tau_i(\theta_0)+\beta_{i+1}h\tau_{i+1}(\theta_0)\leq 2\beta_{i+1}h\tau_{i+1}(\theta_0). \qquad (31)\end{aligned}$$

Note that this last inequality is a statement about $X_t(\theta_0)$ only, and is independent of the construction used. We may drop the argument θ_0.

For fixed i, $P(\tau_{i+1} - \tau_i \leq 2\beta^{i+1}h\tau_{i+1}) = O(h)$ because $\tau_{i+1} - \tau_i$ is exponentially distributed and $1 - e^{-h} = O(h)$. The probability that this holds for fixed i and j, $(j \neq i)$ is $O(h^2)$, by the statistical independence of disjoint spacings. Hence, the probability that this holds for *some* i and j less than a fixed n is $\binom{n}{2}$ times the probability that it holds for a fixed i and j; i.e., it is $O(h^2)$. The case $\tilde{N} = N_T$ is a bit more involved.

If $i \leq N_T$, then $\tau_i \leq T$ so we may replace (29) with $|d\tau_i/d\theta| \leq (B\rho^i/q_*)T$. Since we need to consider only transitions in $(0, T]$, this makes $|\tau_i(\theta_0 + h) - \tau_i(\theta_0)| \leq (B\rho^i/q_*)Th \leq b^iTh$, where $b = \max(1, B/q_*)\rho$. Arguing as in (31), we now need to bound the probability that $|\tau_{i+1}(\theta_0) - \tau_i(\theta_0)| \leq 2b^{i+1}Th$ for more than one i.

Uniformize the process at rate q^*. Let N^* be the number of transitions in $(0, T]$ using uniformization. For $i \leq N^*$, let τ_i^* denote the epoch of the ith such transition, and let $\tau_{N^*+1}^* = T$ (this is conservative because $\tau_{N_T+1} > T$). We view the original (nonhull) transition epochs at a subset of the τ_i^*.

Since all transitions are constrained to occur in $(0, T]$, for $i \leq N^*$ we may replace the bound $b^{i+1}Th$ with b^{N^*+1}, and consider

$$P(\text{for some } i, j \leq N^*,\ \tau_{i+1}^* - \tau_i^* \leq 2b^{N^*+1}Th,\ \tau_{j+1}^* - \tau_j^* \leq 2b^{N^*+1}Th). \qquad (32)$$

Given N^*, the points corresponding to $\{\tau_i^*, i \leq \tilde{N}\}$ are uniformly distributed on $(0, T]$. For fixed, distinct $i, j \leq N^*$ and any $t_i, t_j > 0$,

$$P(\tau_{i+1}^* - \tau_i^* \leq t_i \text{ and } \tau_{j+1}^* - \tau_j^* \leq t_j \mid N^*) \leq \left(\frac{t_i}{T}\right)\left(\frac{t_j}{T}\right).$$

Hence,

$$P(\tau_{i+1}^* - \tau_i^* \leq 2b^{N^*+1}Th \text{ and } \tau_{j+1}^* - \tau_j^* \leq 2b^{N^*+1}Th \mid N^*) \leq (2b^{N^*+1}h)^2.$$

Since there are $\binom{N^*}{2} < N^{*2}$ ways of choosing i and j,

$$P(\text{for some } i, j \leq N^*, \tau_{i+1}^* - \tau_i^* \leq 2b^{N^*+1}Th, \tau_{j+1}^* - \tau_j^* \leq 2b^{N^*+1}Th \mid N^*) \leq (N^* \cdot 2b^{N^*+1}h)^2.$$

Unconditioning, we get $4h^2E[(N^*b^{N^*+1})^2]$. Since N^* has a Poisson distribution, the expectation is finite, and the product is $O(h^2)$.

Proof of Proposition 3. Define, for $j \in C_n$,

$$R_n(j) = \tau_n + c_n(j)/q_n(j); \qquad (33)$$

then, since $c_n(j)$ is the time remaining on clock j at τ_n, $R_n(j)$ is the time at which clock j is *scheduled* to run out, as of τ_n. By definition, J_n runs out at τ_n, so $c_n(J) = 0$; hence, $\tau_n = R_n(J_n)$, and showing that $\tau_n' = D_n(J_n)$ reduces to showing that $R_n'(J_n) = D_n(J_n)$. In fact, we will show that for all j

$$R_n'(j) = D_n(j) - \frac{c_n(j)}{q_n(j)^2}\, q_n'(j), \qquad (34)$$

provided that $j \in C_n$. This implies $R_n'(J_n) = D_n(J_n)$ because $c_n(J_n) = 0$. Also, since J_{n+1} runs out at τ_{n+1},

$$c_n(J_{n+1})/q_n(J_{n+1}) = \tau_{n+1} - \tau_n; \qquad (35)$$

Hence, once we have shown that $R_n'(J_n) = D_n(J_n)$, for all n, (34) will imply that $D_{n+1}(J_{n+1}) = D_n(J_{n+1}) - q_n'(J_{n+1})[\tau_{n+1} - \tau_n]/q_n(J_{n+1})$, which explains (25).

To prove (34), first notice that it holds at $n = 0$ because $D_0(j) \equiv 0$. Take as induction hypothesis that it holds up to n. If $j \in C_{n+1} \cap C_n$, then solving for $c_n(j)$ in (33) and plugging the answer into (22) we get $c_{n+1}(j) = q_n(j)[R_n(j) - \tau_{n+1}]$. Substituting this back into (33) applied to $R_{n+1}(j)$, we get

$$R_{n+1}(j) = \tau_{n+1} + q_n(j)[R_n(j) - \tau_{n+1}]/q_{n+1}(j).$$

As noted, $\tau_{n+1} = R_n(J_{n+1})$. Making this substitution and differentiating yields

$$R_{n+1}'(j) = \left(\frac{q_n(j)}{q_{n+1}(j)}\right)R_n'(j) + \left(1 - \frac{q_n(j)}{q_{n+1}(j)}\right)R_n'(J_{n+1}) + (R_n(j) - R_n(J_{n+1}))\left(\frac{q_n(j)}{q_{n+1}(j)}\right)'.$$

This expresses $R_{n+1}'(j)$—the derivative of the time of the scheduled consumption of j—as the sum of three factors: The first is the derivative in the previous scheduled time, rescaled to the new rate for clock j.

The second is the contribution of a delay in the consumption of J_{n+1}; if J_{n+1} runs out later, j runs longer at $q_n(j)$ and shorter at $q_{n+1}(j)$. Finally, the time it takes to complete the residual scheduled time $R_n(j) - R_n(J_{n+1}) = R_n(j) - \tau_{n+1}$ depends on the rates $q_n(j)$ and $q_{n+1}(j)$; this is reflected in the last term.

Using the induction hypothesis, i.e., (34), this expression for $R'_{n+1}(j)$ becomes

$$\left(\frac{q_n(j)}{q_{n+1}(j)}\right)(D_n(j) - \frac{c_n(j)}{q_n(j)^2}\, q'_n(j))$$
$$+\left(1-\frac{q_n(j)}{q_{n+1}(j)}\right)\left(D_n(J_{n+1})-\frac{c_n(J_{n+1})}{q_n(J_{n+1})^2}q'_n(J_{n+1})\right)$$
$$+(R_n(j)-R_n(J_{n+1}))\left(\frac{q_n(j)}{q_{n+1}(j)}\right)'. \qquad (36)$$

This is also valid when $j \in C_{n+1} \backslash C_n$ provided that we understand $q_n(j)$ and $c_n(j)/q_n(j)^2$ to be identically zero in this case. Now, the residual scheduled time $R_n(j) - R_n(J_{n+1})$ is just $c_{n+1}(j)/q_n(j)$, so we may make this substitution in (36). Also, $c_n(j)$, the time on clock j at τ_n, is $c_{n+1}(j)$ plus the amount j was reduced during $[\tau_n, \tau_{n+1})$, so we have

$$c_n(j) = q_n(j)[\tau_{n+1} - \tau_n] + c_{n+1}(j).$$

Make this substitution in (36) as well. Finally, substitution for $c_n(J_{n+1})$ according to (35). These substitutions make (36)

$$\left(\frac{q_n(j)}{q_{n+1}(j)}\right)\left[D_n(j) - \frac{\tau_{n+1} - \tau_n}{q_n(j)}\, q'_n(j)\right]$$
$$+\left(1-\frac{q_n(j)}{q_{n+1}(j)}\right)\left[D_n(J_{n+1})-\frac{\tau_{n+1}-\tau_n}{q_n(J_{n+1})}q'_n(J_{n+1})\right]$$
$$-\left(\frac{q_n(j)}{q_{n+1}(j)}\right)\left(\frac{c_{n+1}(j)}{q_n^2(j)}\right)q'_n(j)+\left(\frac{c_{n+1}(j)}{q_n(j)}\right)\left(\frac{q_n(j)}{q_{n+1}(j)}\right)'.$$

Comparison with (25) and (26) reveals that the first two terms combine to give simply, $D_{n+1}(j)$; i.e., we have

$$R'_{n+1}(j)=D_{n+1}(j)-\left(\frac{q_n(j)}{q_{n+1}(j)}\right)\left(\frac{c_{n+1}(j)}{q_n^2(j)}\right)q'_n(j)$$
$$+\left(\frac{c_{n+1}(j)}{q_n(j)}\right)\left(\frac{q_n(j)}{q_{n+1}(j)}\right)'.$$

Expanding the last derivative and cancelling terms we get $R'_{n+1}(j) = D_{n+1}(j) - q'_{n+1}(j)c_{n+1}(j)/q^2_{n+1}(j)$, which is what we needed to show; see (34).

Proof of Theorem 1. Let $\Delta_h L = L(\theta + h) - L(\theta)$ for either $L = L_1$ or L_2. Let $\tilde{h}$ be the infimum over h for which L has a discontinuity in $(\theta - h, \theta + h)$ (which is strictly positive with probability one). We consider separately the two terms in

$$E\left[\frac{\Delta_h L}{h}\right]=E\left[\frac{\Delta_h L}{h}\mathbf{1}\{h<\tilde{h}\}\right]+E\left[\frac{\Delta_h L}{h}\mathbf{1}\{h\geq\tilde{h}\}\right].$$

For the first term, we show that

$$\lim_{h\to 0} E\left[\frac{\Delta_h L}{h}\mathbf{1}\{h<\tilde{h}\}\right]=E\left[\frac{dL}{d\theta}\right]. \qquad (37)$$

Since L is continuous and piecewise differentiable on $(\theta - h, \theta + h)$ when $h \leq \tilde{h}$, the mean value theorem implies

$$\left|\frac{\Delta_h L}{h}\mathbf{1}\{h<\tilde{h}\}\right| \leq \sup_{(\theta-h,\theta+h)} \left|\frac{dL}{d\theta}\right|.$$

Then (37) follows from the dominated convergence theorem if we can show that

$E[\sup_{(\theta-h,\theta+h)} | dL/d\theta |] < \infty$.

Let f^* be the supremum of the (bounded) function $|f|$, and let $\rho = 2[1 + (q^*/q_*)]$. Clearly, $\rho > 1$. Note that

$$\max_{j\in C_1} | D_1(j) | \leq B\left(\frac{\tau_1 - \tau_0}{q_*}\right)$$

and

$$\max_{j\in C_{n+1}} | D_{n+1}(j)|$$
$$\leq \rho\left[\max_{j\in C_n} | D_n(j)| + B\left(\frac{\tau_{n+1} - \tau_n}{q_*}\right)\right].$$

Thus,

$$\left|\frac{d\tau_n}{d\theta}\right| = | D_n(J_n)| \leq \sum_{i=0}^{n-1}\left(\frac{\tau_{i+1} - \tau_i}{q_*}\right)B\rho^{n-i-1}$$
$$\leq B\tau_n\rho^n/q_*. \qquad (38)$$

Using (5), we find that

$$\left|\frac{dL_2}{d\theta}\right| \leq 2f^* \sum_{i=1}^{n}\left|\frac{d\tau_i}{d\theta}\right|$$
$$\leq 2f^*(B/q_*) \sum_{i=1}^{n} \tau_i\rho^i$$
$$\leq 2f^*Bn\tau_n\rho^n/q_*.$$

Since $\sup_{(\theta-h,\theta+h)}\tau_n$ is bounded by $(X_1 + \ldots + X_n)/q_*$, which is integrable, the result follows for L_2. For L_1, we get, in the same way, the bound

$$\left|\frac{dL_1}{d\theta}\right| \leq (2f^*B/q_*)\tau_{N_T}N_T\rho^{N_T} \leq (2f^*B/q_*)TN_T\rho^{N_T}.$$

At each θ, N_T is stochastically bounded by a Poisson random variable with parameter q^*T; hence, so is $\sup_{(\theta-h,\theta+h)}(N_T)$. For a Poisson random variable $\hat{N}$, a simple calculation shows that $E[\hat{N}\rho^{\hat{N}}] < \infty$, which concludes verification of (37).

For the second term, we need to verify that as $h \to 0$, $E[\Delta_h L/h)\mathbf{1}\{h \geq \tilde{h}\}] \to 0$. Using the Cauchy-Schwarz inequality; we get

$$E\left[\left(\frac{\Delta_h L}{h}\mathbf{1}\{h \geq \tilde{h}\}\right)^2\right] \leq E[\Delta_h L^2]\frac{P(\tilde{h} \leq h)}{h^2}.$$

Since $P(\tilde{h} \leq h) = O(h^2)$ (Lemma 2) we need only verify that $E[\Delta_h L^2]$ goes to 0. With probability one, $\Delta_h L \to 0$ as $h \to 0$; and for all θ, $E[L_1^2(\theta)] \leq (f^*T)^2$ and $E[L_2^2(\theta)] \leq 2n(f^*/q_*)^2$ so the dominated convergence theorem yields the result.

10. CONCLUDING REMARKS

The results of this paper show that good derivative estimates are available for a range of functionals of continuous-time Markov chains. Preliminary experience suggests using infinitesimal perturbation analysis if possible and the likelihood ratio method otherwise. In some cases, it is reasonable to use both: for example, if the distribution of the initial state depends on θ, then it should be possible to use likelihood ratios to capture this dependence while perturbation analysis is used to capture that of $Q(\theta)$, assuming **CM** is satisfied. L'Ecuyer (1989) proposes similar combinations of the two methods.

ACKNOWLEDGMENT

Thanks go to Craig Tovey and Adam Rosenberg for several interesting and helpful discussions while they were at Bell Labs. The comments of the referees improved the organization of the paper. Ben Fox provided many valuable, detailed comments which led to several improvements, especially in Sections 5 and 6. This work was carried out while the author was with the Operations Research Department of AT&T Bell Laboratories.

REFERENCES

FOX, B. L. 1990. Generating Markov-Chain Transitions Quickly: I. *ORSA J. Comput.* **2,** 126–135.

FOX, B. L., AND P. W. GLYNN. 1986. Discrete-Time Conversion for Simulating Semi-Markov Processes. *Opns. Res. Letts.* **5,** 191–196.

→ FOX, B. L., AND P. W. GLYNN. 1990. Discrete-Time Conversion for Simulating Finite-Horizon Markov Processes. *SIAM J. Appl. Math.* **50,** 1457–1473.

GLASSERMAN, P. 1988a. Infinitesimal Perturbation Analysis of a Birth and Death Process. *Opns. Res. Letts.* **7,** 43–49.

GLASSERMAN, P. 1988b. Performance Continuity and Differentiability in Monte Carlo Optimization, In *Proceedings of the Winter Simulation Conference,* The Society for Computer Simulation, San Diego, Calif., 518–524.

GLASSERMAN, P. 1990. Discrete-Time "Inversion" and Derivative Estimation for Markov Chains. *Opns. Res. Letts.* **9,** 305–313.

→ GLASSERMAN, P. 1991. Structural Conditions for Perturbation Analysis Derivative Estimation: Finite Time Performance Indices. *Opns. Res.* **39,** 724–738.

GLYNN, P. W. 1986. Stochastic Approximation for Monte Carlo Optimization. In *Proceedings of the Winter Simulation Conference,* The Society for Computer Simulation, San Diego, Calif., 365–364.

→ GLYNN, P. W., AND D. L. IGLEHART. 1989. Importance Sampling for Stochastic Simulations. *Mgmt. Sci.* **35,** 1367–1392.

GLYNN, P. W., AND W. WHITT. 1989. The Efficiency of Simulation Estimators. AT&T Bell Laboratories Technical Report, Murray Hill, N.J.

→ HEIDELBERGER, P., X. R. CAO, M. A. ZAZANIS AND R. SURI. 1988. Convergence Properties of Infinitesimal Perturbation Analysis Estimates. *Mgmt. Sci.* **34,** 1281–1302.

HO, Y. C. 1987. Performance Evaluation and Perturbation Analysis of Discrete Event Dynamical Systems, Perspectives and Open Problems. *IEEE Trans. Autom. Control,* **AC-32,** 563–572.

HO, Y. C., AND S. LI. 1988. Extensions of Perturbation Analysis of Discrete-Event Dynamic Systems. *IEEE Trans. Autom. Control* **AC-33,** 427–438.

→ HORDIJK, A., D. L. IGLEHART AND R. SCHASSBERGER. 1976. Discrete Time Methods for Simulating Continuous Time Markov Chains. *Adv. Appl. Prob.* **8,** 772–788.

L'ECUYER, P. 1990. A Unified View of the IPA, SF and LR Gradient Estimation Techniques. *Mgmt. Sci.* **36,** 1364–1383.

NEUTS, M. F. 1981. *Matrix Geometric Solutions to Stochastic Models.* The Johns Hopkins University Press, Baltimore.

→ REIMAN, M. I., AND A. WEISS. 1989. Sensitivity Analysis for Simulations via Likelihood Ratios. *Opns. Res.* **37,** 830–844.

→ RUBINSTEIN, R. 1989. Sensitivity Analysis and Performance, Extrapolation for Computer Simulation Models. *Opns. Res.* **37,** 72–81.

SURI, R. 1989. Perturbation Analysis: The State of the Art and Research Issues Explained via the GI/G/1 Queue. *Proc. IEEE* **77,** 114–137.