

Exact Simulation of Continuous Time Markov Jump Processes with Anticorrelated Variance Reduced Monte Carlo Estimation

Peter A. Maginnis, Matthew West and Geir E. Dullerud

Abstract— We provide an exact, continuous time extension to previous work in anticorrelated stochastic process simulation that was performed in an approximate, discrete time setting. These methods reduce the variance of continuous time Monte Carlo for Markov jump process systems. We rigorously construct antithetic Poisson processes and analytically prove the negative correlation between pairs. We then show how these anticorrelated Poisson processes can be used to drive Markov jump processes via a random time change representation. Finally, we provide a sufficient condition for variance reduction in the jump process context as well as demonstrate a simple example.

I. INTRODUCTION

In the broad study of stochastic systems, simulation is an increasing necessity. With the proliferation of nonlinear, non-Gaussian applications of interest in filtering and estimation, for example, often stochastic simulation is the only tractable option. The cost of such computations grows dramatically with the scale, complexity or desired accuracy of the simulation, and practical, provable methods to increase efficiency of simulation are often indispensable. A common problem is the reduction of variance of Monte Carlo estimates of system features. In previous work [8], [7], we have defined and proved several anticorrelated variance reduction techniques for the simulation and estimation of Markov jump processes using a discrete-time tau leaping approximation. In this work, we extend those efforts to an exact, continuous-time domain. Specifically, we rigorously construct principle-based simulation schemes that produce anticorrelated ensembles of continuous time Markov jump process paths that can be used to construct reduced variance mean estimators. The crux of the approach is the random time-change (RTC) representation of a Markov jump process due to Kurtz [5]:

$$X(t) = x_0 + \sum_{i=1}^I Y^i \left(\int_{s=0}^t \rho^i(s, X(s)) ds \right) \zeta^i, \quad (1)$$

driven by independent, unit-rate Poisson processes Y^i . This model description is extremely general, and can be used to construct virtually any discrete-space Markov process. Inspired by classical anticorrelated variance reduction techniques (primarily here pairs of antithetic variates), we carefully construct pairs of realizations $(Y^{i,(1)}, Y^{i,(2)})$ of the Poisson process that are pairwise negatively correlated. Using I independent realizations of these anticorrelated pairs of Poisson processes, we may construct anticorrelated realizations $(X^{(1)}, X^{(2)})$ of the Markov jump process of interest, producing an ensemble from which to construct variance-reduced estimates of the mean behavior of the process.

The variance reduction approach of producing anticorrelated realizations of complicated stochastic processes by manipulating the elemental processes that drive them is already in use in many contexts. The most common is probably the antithetic simulation of Brownian motion to produce continuous time Markov processes, where the construction of perfect antithetic pairs is trivial due to symmetry. Producing valid realizations of the Poisson process that are significantly negatively correlated requires somewhat more effort, and the inherent asymmetry of the process does not readily demonstrate optimality of any particular approach. The Markov jump processes that they drive, however, are of tremendous theoretical and practical interest, for example in the study of chemical reaction networks [9], gene regulation networks [2], and atmospheric aerosol simulation [10]. In each of these areas, variance reduction techniques are of great benefit to reduce the cost of stochastic simulation. The manipulation of the random process inputs of the RTC as a variance reduction approach has precedents as well, for example the common reaction path method of Rathinam, et al. [9] and the strong coupling of Poisson process inputs by Anderson [1], both in service of parameter sensitivity analysis. Furthermore, the use of related anticorrelated ensemble techniques in the context of the discrete-time tau-leaping approximation of the Markov jump process is the subject of [8]. The primary contribution of this paper is the construction of antithetic Poisson processes and their use in the RTC for the purpose of exact simulation and variance reduced mean estimation of continuous time processes.

The sequel is organized as follows. First, we propose several techniques to produce anticorrelated unit rate Poisson processes using any existing anticorrelated technique, though here only explicitly considering antithetic pairs. For each method, we provide analytical and numerical support, and illustrate the usefulness and challenges of each. Next we show how to use these methods in support of simulation of Markov jump processes, as well as prove a sufficient condition for variance reduction in this setting. Finally, we provide a simple numerical example illustrating the potential effectiveness of the method.

II. ANTICORRELATED CONTINUOUS-TIME POISSON PROCESSES

We begin by defining several algorithms to exactly simulate Poisson processes, while ensuring that they be as negatively correlated as possible. Such an anticorrelated ensemble of realizations could then be used to construct reduced variance mean estimates of the process, and, moreover, be

used to produce anticorrelated ensembles of general Markov jump process paths via the RTC.

Consider a unit-rate Poisson process $Y(t)$ over a finite time interval $[0, T]$. We define sampling algorithms to produce anticorrelated sample paths of this process in order to create variance reduced pathwise mean estimators. In order to produce pathwise anticorrelated processes, we require in addition a method α of random variable simulation to produce Poisson variables $\{X^{(j)}\}_{j=1}^M$, $X_j \stackrel{\alpha}{\sim} \text{Pois}(\lambda)$, an ensemble of pairwise negatively correlated Poisson random variables such that $\text{Cov}(X^{(j_1)}, X^{(j_2)}) \leq 0$. Without loss of generality, we restrict our analysis to an antithetic method α producing pairwise antithetic variables $X^{(1)} := F_\lambda^{-1}(U)$ and $X^{(2)} := F_\lambda^{-1}(1-U)$, where $U \sim \text{Unif}(0, 1)$ and F_λ^{-1} is the formal inverse of the CDF of a Poisson random variable with parameter λ . This is a well-known, classical technique [11], but its use in this specific context is covered in detail in [6] and [8].

Our objective is to produce pairs of Poisson process sample paths $(Y^{(1)}, Y^{(2)})$ such that, for every $t \in [0, T]$, $\text{Cov}(Y^{(1)}(t), Y^{(2)}(t)) \leq 0$. The motivation for this is that an estimator of the mean behavior $\mathbb{E}[Y(t)]$ of the process, say $\Psi(t) = \frac{1}{2}(Y^{(1)}(t) + Y^{(2)}(t))$, will have variance given by

$$\begin{aligned} \text{Var}(\Psi(t)) &= \frac{1}{2} \text{Var}(Y(t)) + \frac{1}{2} \text{Cov}(Y^{(1)}(t), Y^{(2)}(t)) \quad (2) \\ &\leq \frac{1}{2} \text{Var}(Y(t)), \quad (3) \end{aligned}$$

where the last term would be the variance of an estimator constructed using independent Poisson process paths.

The first approach to producing such ensembles is the endpoint approach. One way to simulate a Poisson process is to sample the value of the process (a Poisson random variable denoting the number of jumps occurring by that time) at some endpoint T of time and then sample when all of those jumps occurred in $[0, T]$ (which will each have $\text{Unif}(0, T)$ distribution). The idea of the endpoint method is to instead simulate two Poisson processes simultaneously by using an antithetic pair of Poisson random variates as their values at time T . Thus by construction, these two values, $Y_E^{(1)}(T)$ and $Y_E^{(2)}(T)$, will be negatively correlated. We simulate the jump times as in the naive i.i.d. case. The endpoint method is summarized in Algorithm 1. A pair of antithetic sample

Algorithm 1 Endpoint Method

- 1) Using method α , simulate anticorrelated Poisson random variables $Y_E^{(j)}(T) \stackrel{\alpha}{\sim} \text{Pois}(T)$
 - 2) Simulate the corresponding jump times $t_i^{(j)}$ i.i.d. $\text{Unif}(0, T)$, $i \in \{1, \dots, Y_E^{(j)}(T)\}$ for each path $j \in \{1, \dots, M\}$
-

paths simulated using the endpoint method are shown in Figure 1. Now, while the two paths are explicitly negatively correlated at time T , they are also negatively correlated throughout the time interval, and thus, a pathwise mean

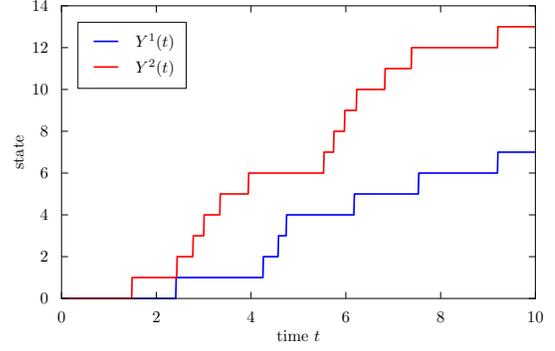


Fig. 1. Two anticorrelated sample paths of a unit rate Poisson process simulated using the endpoint antithetic pathwise sampling technique. In this case $Y_E^{(1)}(10), Y_E^{(2)}(10) \stackrel{\text{anti.}}{\sim} \text{Pois}(10)$ are antithetically sampled.

estimator constructed using these samples will have reduced variance relative to a naive estimator throughout the time interval, as shown in Lemma 2. To prove this result, we first prove a very useful Lemma that computes the covariance at interior points in terms of covariances at endpoints of certain intervals.

Lemma 1: Let $Y^{(1)}(t)$ and $Y^{(2)}(t)$, $t \in [0, T]$ be two (possibly correlated) unit rate Poisson processes. For any $T_1 < T_2 \in [0, T]$, define $\mathcal{G}(T_1, T_2) := \sigma\{Y^{(1)}(T_1), Y^{(1)}(T_2), Y^{(2)}(T_1), Y^{(2)}(T_2)\}$, the σ -algebra defined by these four random variables. If, for any $t \in (T_1, T_2)$, conditioned on $\mathcal{G}(T_1, T_2)$, $Y^{(1)}(t)$ and $Y^{(2)}(t)$ are independent, then for every $t \in [T_1, T_2]$:

$$\begin{aligned} \text{Cov}(Y^{(1)}(t), Y^{(2)}(t)) &= \text{Cov}(Y^{(1)}(T_1), Y^{(2)}(T_1)) \\ &+ \frac{(t - T_1)^2}{(T_2 - T_1)^2} \text{Cov}(N^{(1)}(T_1, T_2), N^{(2)}(T_1, T_2)) \quad (4) \end{aligned}$$

where $N^{(j)}(T_1, T_2) := Y^{(j)}(T_2) - Y^{(j)}(T_1)$ is the increment of the process. Alternatively:

$$\begin{aligned} \text{Cov}(Y^{(1)}(t), Y^{(2)}(t)) &= \frac{T_2 - t}{(T_2 - T_1)^2} (T_2 + t - 2T_1) \text{Cov}(Y^{(1)}(T_1), Y^{(2)}(T_1)) \\ &+ \frac{(t - T_1)^2}{(T_2 - T_1)^2} \text{Cov}(Y^{(1)}(T_2), Y^{(2)}(T_2)) \quad (5) \end{aligned}$$

Proof: For the sake of brevity, we prove only (4). The proof of (5) is similar but more algebraically tedious.

$$\begin{aligned} \mathbb{E}[Y^{(1)}(t)Y^{(2)}(t)] &= \mathbb{E}\left[\mathbb{E}\left[Y^{(1)}(t)Y^{(2)}(t) \middle| \mathcal{G}\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[Y^{(1)}(t) \middle| \mathcal{G}\right] \mathbb{E}\left[Y^{(2)}(t) \middle| \mathcal{G}\right]\right] \\ &= \mathbb{E}\left[\left[Y^{(1)}(T_1) + \frac{t - T_1}{T_2 - T_1} N^{(1)}(T_1, T_2)\right] \right. \\ &\quad \cdot \left. \left[Y^{(2)}(T_1) + \frac{t - T_1}{T_2 - T_1} N^{(2)}(T_1, T_2)\right]\right] \\ &= \mathbb{E}\left[Y^{(1)}(T_1)Y^{(2)}(T_1)\right] + 2T_1(t - T_1) \\ &\quad + \frac{(t - T_1)^2}{(T_2 - T_1)^2} \mathbb{E}\left[N^{(1)}(T_1, T_2)N^{(2)}(T_1, T_2)\right]. \end{aligned}$$

Note that

$$t^2 = T_1^2 + 2T_1(t - T_1) + \frac{(t - T_1)^2}{(T_2 - T_1)^2}(T_2 - T_1)^2, \quad (6)$$

so that

$$\text{Cov}(Y^{(1)}(t), Y^{(2)}(t)) = \mathbb{E} \left[Y^{(1)}(T_1)Y^{(2)}(T_1) \right] - t^2 \quad (7)$$

$$= \mathbb{E} \left[Y^{(1)}(T_1)Y^{(2)}(T_1) \right] - T_1^2 \\ + \frac{(t - T_1)^2}{(T_2 - T_1)^2} \left(\mathbb{E} \left[N^{(1)}(T_1, T_2)N^{(2)}(T_1, T_2) \right] - (T_2 - T_1)^2 \right) \quad (8)$$

$$= \text{Cov}(Y^{(1)}(T_1), Y^{(2)}(T_1)) \\ + \frac{(t - T_1)^2}{(T_2 - T_1)^2} \text{Cov}(N^{(1)}(T_1, T_2), N^{(2)}(T_1, T_2)), \quad (9)$$

and the claim holds. \blacksquare

This Lemma may now be immediately used to prove that there is negative correlation between endpoint antithetic Poisson process pairs at every $t \in [0, T]$, not just the endpoint.

Lemma 2: Suppose $Y_E^{(1)}(t), Y_E^{(2)}(t)$ are simulated using Algorithm 1 for $t \in [0, T]$ and some anticorrelated sampling method α . Then $\text{Cov}(Y_E^{(1)}(t), Y_E^{(2)}(t)) \leq 0$, for every $t \in [0, T]$.

Proof: We proceed by application of Lemma 1. First, note that, by construction $\text{Cov}(Y_E^{(1)}(T), Y_E^{(2)}(T)) \leq 0$ since these are precisely the antithetically sampled Poisson random variables. Further, note that, conditioned on $Y^{(1)}(T)$ and $Y^{(2)}(T)$, $Y^{(1)}(t)$ and $Y^{(2)}(t)$ are independent for every $t \in (0, T)$. Thus, by (4),

$$\text{Cov}(Y_E^{(1)}(t), Y_E^{(2)}(t)) = \frac{t^2}{T^2} \text{Cov}(N_E^{(1)}(0, T), N_E^{(2)}(0, T)) \quad (10)$$

$$= \frac{t^2}{T^2} \text{Cov}(Y_E^{(1)}(T), Y_E^{(2)}(T)) \leq 0,$$

for every $t \in [0, T]$. \blacksquare

Thus, at time T , the paths will have the same negative correlation as is produced by the random variable method, and this negative correlation will be felt by the interior points with a quadratic relationship. As a direct consequence of this result, unbiased pathwise mean estimators constructed using these anticorrelated samples will have reduced variance.

Corollary 1: Suppose $\Psi(t) := \frac{1}{2}(Y^{(1)}(t) + Y^{(2)}(t))$ and $\Psi_E(t) := \frac{1}{2}(Y_E^{(1)}(t) + Y_E^{(2)}(t))$ are two unbiased mean estimators of $\mathbb{E}[Y(t)]$ constructed using sample paths drawn i.i.d. and via the antithetic endpoint technique, respectively. Then $\text{Var}(\Psi(t)) \geq \text{Var}(\Psi_E(t)) = \text{Var}(\Psi(t)) + \frac{1}{2} \text{Cov}(Y_E^{(1)}(t), Y_E^{(2)}(t))$.

Figure 2 illustrates this fact by plotting the variance of two 4-sample mean estimators, one using the traditional Monte Carlo estimation scheme (4 i.i.d. sample paths) and one using two pairs of endpoint antithetically sampled Poisson process paths versus time. Here, we observe a dramatic reduction of variance at time $t = 10$, and some gains for smaller time.

We may improve performance in the interior of the time interval via an obvious extension, the Concatenation technique. By exploiting the independent increments property of

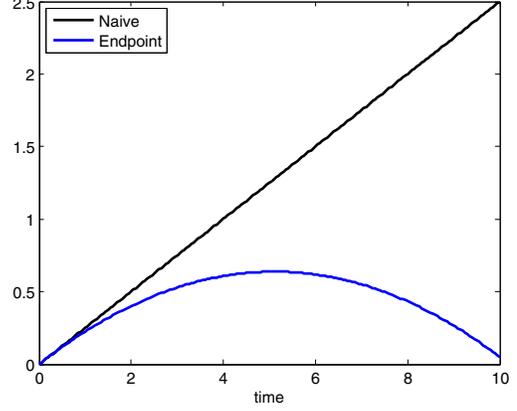


Fig. 2. Comparison variances of a 4-sample unit-rate Poisson process endpoint antithetic mean estimator $\Psi_E^4(t)$ to the naive (i.i.d.) 4-point mean estimator $\Psi^4(t)$.

Poisson processes, we split the interval $[0, T]$ into smaller sub-intervals and simulate each Poisson process increment using the endpoint technique, as shown in Alg. 2. A pair

Algorithm 2 Concatenation Method

- 1) Divide the interval $[0, T]$ into K equal length sub-intervals
 - 2) Using method α , simulate anticorrelated Poisson random variable increments $N_C^{(j)}(T^{\frac{k-1}{K}}, T^{\frac{k}{K}}) := Y_C^{(j)}(T^{\frac{k}{K}}) - Y_C^{(j)}(T^{\frac{k-1}{K}}) \stackrel{\alpha}{\sim} \text{Pois}(\frac{T}{K})$ for $k \in \{1, \dots, K\}$
 - 3) Simulate the corresponding jump times $t_i^{(j)} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(T^{\frac{k-1}{K}}, T^{\frac{k}{K}})$, $i \in \{Y_C^{(j)}(T^{\frac{k-1}{K}}) + 1, \dots, Y_C^{(j)}(T^{\frac{k}{K}})\}$, $k \in \{1, \dots, K\}$ for each path $j \in \{1, \dots, M\}$
-

of sample paths simulated using the concatenated antithetic sampling technique for $K = 2$ and $T = 10$ are shown in Fig. 3. As shown in Fig. 4, by splitting the time interval in half, we achieve another large reduction in the variance of a mean estimator at the concatenation point where an antithetic pair of Poisson random variables are used. Note also that, in this case there is a small sacrifice in performance at the endpoint $t = 10$. This is due to the independent increments property implying that the covariances of sums of increments are sums of covariances of increments, resulting in an upward drift of the estimator variance at the endpoint. As in Corollary 1, the following Lemma is equivalent to variance reduction of a mean estimator using these sample paths.

Lemma 3: Suppose $Y_C^{(1)}(t), Y_C^{(2)}(t)$ are simulated using Algorithm 2 for $t \in [0, T]$ and some anticorrelated sampling method α . Then $\text{Cov}(Y_C^{(1)}(t), Y_C^{(2)}(t)) \leq 0$, for every $t \in [0, T]$.

Proof: First, note that each increment $N_C^{(j)}(T^{\frac{k-1}{K}}, T^{\frac{k}{K}}) = Y_C^{(j)}(T^{\frac{k}{K}}) - Y_C^{(j)}(T^{\frac{k-1}{K}})$, $k \in$

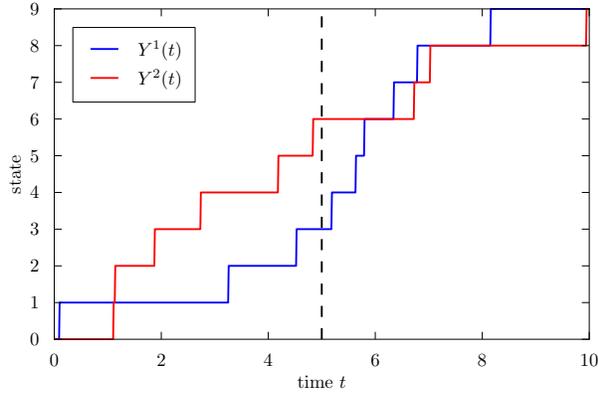


Fig. 3. Pair of anticorrelated unit-rate Poisson process sample paths simulated using concatenated antithetic pathwise sampling. Note here that $Y_C^{(1)}(5), Y_C^{(2)}(5) \stackrel{\text{anti}}{\sim} \text{Pois}(5)$ are independent of $(Y_C^{(1)}(10) - Y_C^{(1)}(5))$ and $(Y_C^{(2)}(10) - Y_C^{(2)}(5)) \stackrel{\text{anti}}{\sim} \text{Pois}(5)$.

$\{1, \dots, K\}$ satisfies

$$\text{Cov} \left(N_C^{(1)} \left(T \frac{k-1}{K}, T \frac{k}{K} \right), N_C^{(2)} \left(T \frac{k-1}{K}, T \frac{k}{K} \right) \right) \leq 0 \quad (11)$$

since these are each antithetically sampled random variables. Next, it is easy to see that the conditional independence requirement of Lemma 1 is satisfied by the process within each small interval, since $Y_C^{(1)}(t), Y_C^{(2)}(t)$ are independent conditioned on $\mathcal{G}(T \frac{k-1}{K}, T \frac{k}{K})$ for each $t \in (T \frac{k-1}{K}, T \frac{k}{K})$ by construction.

Thus for the leftmost interval, by (4), we have that

$$\begin{aligned} & \text{Cov}(Y_C^{(1)}(t), Y_C^{(2)}(t)) \\ &= \frac{t^2}{\left(\frac{T}{K}\right)^2} \text{Cov} \left(N_C^{(1)} \left(0, \frac{T}{K} \right), N_C^{(2)} \left(0, \frac{T}{K} \right) \right) \end{aligned} \quad (12)$$

for every $t \in [0, \frac{T}{K}]$. By the same reasoning, for $t \in [\frac{T}{K}, \frac{2T}{K}]$,

$$\begin{aligned} & \text{Cov} \left(Y_C^{(1)}(t), Y_C^{(2)}(t) \right) \\ &= \text{Cov} \left(Y_C^{(1)} \left(\frac{T}{K} \right), Y_C^{(2)} \left(\frac{T}{K} \right) \right) \\ &+ \frac{(t - \frac{T}{K})^2}{\left(\frac{T}{K}\right)^2} \text{Cov} \left(N_C^{(1)} \left(\frac{T}{K}, \frac{2T}{K} \right), N_C^{(2)} \left(\frac{T}{K}, \frac{2T}{K} \right) \right) \quad (13) \\ &= \text{Cov} \left(N_C^{(1)} \left(0, \frac{T}{K} \right), N_C^{(2)} \left(0, \frac{T}{K} \right) \right) \\ &+ \frac{(t - \frac{T}{K})^2}{\left(\frac{T}{K}\right)^2} \text{Cov} \left(N_C^{(1)} \left(\frac{T}{K}, \frac{2T}{K} \right), N_C^{(2)} \left(\frac{T}{K}, \frac{2T}{K} \right) \right). \quad (14) \end{aligned}$$

By continuing this process from left to right, it is easy to

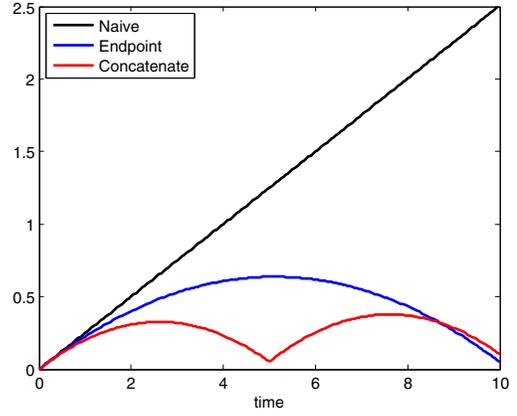


Fig. 4. Comparison of variance of a 4-sample mean estimator $\Psi_C^4(t)$ of a unit-rate Poisson process constructed using $K = 2$ step concatenated antithetic sample paths to variances of a 4-sample endpoint antithetic mean estimator $\Psi_E^4(t)$ and the naive 4-sample mean estimator $\Psi^4(t)$.

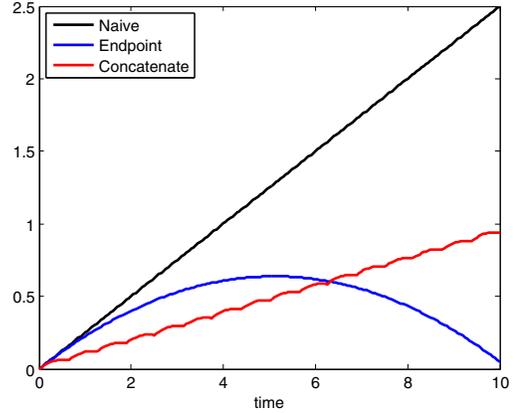


Fig. 5. Comparison of variances of a 4-sample mean estimator $\Psi_C^4(t)$ of a unit-rate Poisson process constructed using $K = 16$ step concatenated antithetic sample paths to a 4-sample endpoint antithetic mean estimator $\Psi_E^4(t)$ and the naive 4-sample mean estimator $\Psi^4(t)$.

see that, for any $t \in [0, T]$.

$$\begin{aligned} & \text{Cov} \left(Y_C^{(1)}(t), Y_C^{(2)}(t) \right) \\ &= \sum_{k=1}^{k^*} \text{Cov} \left(N_C^{(1)} \left(\frac{(k-1)T}{K}, \frac{kT}{K} \right), N_C^{(2)} \left(\frac{(k-1)T}{K}, \frac{kT}{K} \right) \right) \\ &+ \frac{(t - \frac{k^*T}{K})^2}{\left(\frac{T}{K}\right)^2} \\ &\cdot \text{Cov} \left(N_C^{(1)} \left(\frac{k^*T}{K}, \frac{(k^*+1)T}{K} \right), N_C^{(2)} \left(\frac{k^*T}{K}, \frac{(k^*+1)T}{K} \right) \right), \quad (15) \end{aligned}$$

where $k^* := \max\{k \in \mathbb{N} : \frac{kT}{K} \leq t\} = \lfloor \frac{tK}{T} \rfloor$. Since each of these covariances is negative, the claim holds. ■

A direct consequence of this fact is that concatenation may not perform better than the endpoint technique as the number of concatenation steps K increases. Fig. 5 illustrates

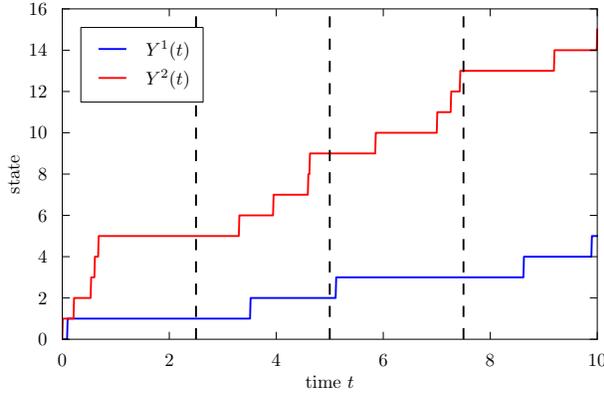


Fig. 6. An anticorrelated pair of sample paths produced using $L = 2$ (comparable to a 4-step concatenated path). Note here that $Y_B^{(j)}(10) \stackrel{\text{anti}}{\sim} \text{Pois}(10)$ is sampled first, then $Y_B^{(j)}(5)|Y_B^{(j)}(10) \stackrel{\text{anti}}{\sim} \text{Binom}(Y_B^{(j)}(10), 1/2)$, and finally $Y_B^{(j)}(2.5)|Y_B^{(j)}(5) \stackrel{\text{anti}}{\sim} \text{Binom}(Y_B^{(j)}(5), 1/2)$ and $[Y_B^{(j)}(7.5) - Y_B^{(j)}(5)]\{Y_B^{(j)}(5), Y_B^{(j)}(10)\} \stackrel{\text{anti}}{\sim} \text{Binom}(5, 1/2)$.

this problem by showing the drift in variance that occurs when 16 concatenation steps are used.

A more ideal mean estimator is one who's endpoint performance matches that of the endpoint algorithm but that further improves performance away from the endpoint T . We achieve such an algorithm by again exploiting the fact that, conditioned on its future value, a Poisson process has binomial distribution. This algorithm, referred to here as binomial midpoint simulation method, is defined in Algorithm 3. The key discovery is that, once we sample the values of the process at the endpoints of an interval, we may conditionally sample the midpoint of that interval using a binomial distribution. When we perform this sampling, we use the anticorrelated method α , reducing the variance of an estimator at this midpoint. We may then recursively sample the midpoints of these new intervals, repeatedly bijecting until we have 2^L steps. We then finish the simulation of the Poisson process by sampling i.i.d. uniform jump times, similar to the concatenation method.

Algorithm 3 Binomial Midpoint Method

- 1) Divide the interval $[0, T]$ into 2^L subintervals
 - 2) Simulate $Y_B^{(j)}(T) \stackrel{\alpha}{\sim} \text{Pois}(T)$ as in Alg. 1
 - 3) Simulate $Y_B^{(j)}(\frac{T}{2})|Y_B^{(j)}(T) \stackrel{\alpha}{\sim} \text{Binom}(Y_B^{(j)}(T), \frac{1}{2})$
 - 4) Fix $\ell = 2, \dots, L$. Then, for each $k \in \{1, 3, 5, \dots, 2^\ell - 1\}$ simulate $N_B^{(j)}(T \frac{k-1}{2^\ell}, T \frac{k}{2^\ell})|\{Y_B^{(j)}(T \frac{k+1}{2^\ell}), Y_B^{(j)}(T \frac{k-1}{2^\ell})\} \stackrel{\alpha}{\sim} \text{Binom}(Y_B^{(j)}(T \frac{k+1}{2^\ell}) - Y_B^{(j)}(T \frac{k-1}{2^\ell}), \frac{1}{2})$
 - 5) Simulate jump times $t_i^{(j)} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(T \frac{k-1}{2^L}, T \frac{k}{2^L})$, $i \in \{Y_B^{(j)}(T \frac{k-1}{2^L}) + 1, \dots, Y_B^{(j)}(T \frac{k}{2^L})\}$, $k \in \{1, \dots, 2^L\}$ for each path $j \in \{1, \dots, M\}$
-

A pair of binomial antithetic sample paths for $L = 2$ is shown in Fig. 6. While dividing the interval by powers of 2 is

not strictly necessary, dyadic intervals are chosen to produce binomial distributions with $p = \frac{1}{2}$, since antithetic sampling is more effective for symmetric distributions. The algorithm performs better as the number of midpoints increases, until reaching a saturation point when midpoints become nearly constant, i.e. when the expected number of transitions in an interval is small. As in Corollary 1, the following Lemma is equivalent to variance reduction of a mean estimator using these sample paths.

Lemma 4: Suppose $Y_B^{(1)}(t), Y_B^{(2)}(t)$ are simulated using Algorithm 3 for $t \in [0, T]$ and some anticorrelated sampling method α . Then $\text{Cov}(Y_B^{(1)}(t), Y_B^{(2)}(t)) \leq 0$, for every $t \in [0, T]$.

Proof: As in the endpoint method case, $\text{Cov}(Y_B^{(1)}(T), Y_B^{(2)}(T)) \leq 0$ as this is an antithetic pair of Poisson samples. For any interior interval boundary point $t^* = \frac{k}{2^\ell}$, $k \in \{1, 3, \dots, 2^\ell - 1\}$, $Y_B^{(1)}(t^*)$ and $Y_B^{(2)}(t^*)$ are a conditionally sampled antithetic pair, that is

$$\text{Cov}\left(Y_B^{(1)}(t^*), Y_B^{(2)}(t^*) \middle| \mathcal{G}\left(t^* - \frac{1}{2^\ell}, t^* + \frac{1}{2^\ell}\right)\right) \leq 0, \quad (16)$$

where the left hand side is a $\mathcal{G}(t^* - \frac{1}{2^\ell}, t^* + \frac{1}{2^\ell})$ -measurable random variable. Furthermore,

$$\begin{aligned} & \mathbb{E}\left[Y_B^{(j)}(t^*) \middle| \mathcal{G}\left(t^* - \frac{1}{2^\ell}, t^* + \frac{1}{2^\ell}\right)\right] \\ &= \frac{1}{2}\left(Y_B^{(j)}\left(t^* - \frac{1}{2^\ell}\right) + Y_B^{(j)}\left(t^* + \frac{1}{2^\ell}\right)\right) \end{aligned} \quad (17)$$

so that

$$\begin{aligned} & \text{Cov}\left(\mathbb{E}\left[Y_B^{(1)}(t^*) \middle| \mathcal{G}\right], \mathbb{E}\left[Y_B^{(2)}(t^*) \middle| \mathcal{G}\right]\right) \\ &= \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_-^*), Y_B^{(2)}(t_-^*)) \\ &+ \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_-^*), Y_B^{(2)}(t_+^*)) \\ &+ \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_+^*), Y_B^{(2)}(t_-^*)) \\ &+ \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_+^*), Y_B^{(2)}(t_+^*)), \end{aligned} \quad (18)$$

where t_+^* and t_-^* denote the ‘‘parents’’ of t^* , $t^* - \frac{1}{2^\ell} = \frac{k-1}{2^{\ell-1}}$ and $t^* + \frac{1}{2^\ell} = \frac{k}{2^{\ell-1}}$, respectively. Thus, by the law of total covariance

$$\begin{aligned} & \text{Cov}\left(Y_B^{(1)}(t^*), Y_B^{(2)}(t^*)\right) \\ &= \mathbb{E}\left[\text{Cov}\left(Y_B^{(1)}(t^*), Y_B^{(2)}(t^*) \middle| \mathcal{G}\left(t^* - \frac{1}{2^\ell}, t^* + \frac{1}{2^\ell}\right)\right)\right] \\ &+ \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_-^*), Y_B^{(2)}(t_-^*)) \\ &+ \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_-^*), Y_B^{(2)}(t_+^*)) \\ &+ \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_+^*), Y_B^{(2)}(t_-^*)) \\ &+ \frac{1}{4}\text{Cov}(Y_B^{(1)}(t_+^*), Y_B^{(2)}(t_+^*)). \end{aligned} \quad (20)$$

Note that the first term is negative, and the argument repeats for each of the other terms, rewriting them as a linear

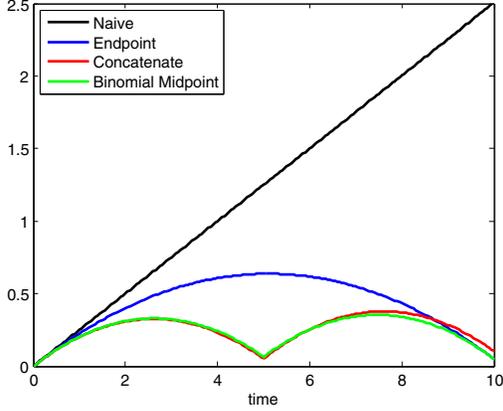


Fig. 7. Comparison of variance of a 4-sample unit-rate Poisson process mean estimator $\Psi_B^4(t)$ constructed using 2-step binomial midpoint antithetic sample paths to the variances of a 2-step concatenated antithetic $\Psi_C^4(t)$, endpoint antithetic $\Psi_E^4(t)$ and naive $\Psi^4(t)$ equivalents.

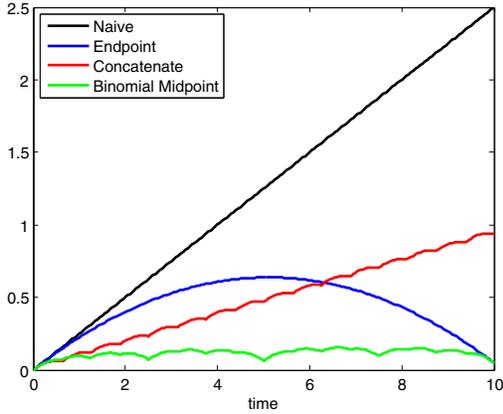


Fig. 8. Comparison of variance of a 4-sample unit-rate Poisson process mean estimator $\Psi_B^4(t)$ constructed using 16-step binomial midpoint antithetic sample paths to the 2-step concatenated antithetic $\Psi_C^4(t)$, endpoint $\Psi_E^4(t)$ antithetic and naive $\Psi^4(t)$ equivalents.

combination of some negative value plus a covariance of $Y^{(1)}$ and $Y^{(2)}$ evaluated at their “parent” times. These recursions will terminate with $\text{Cov}(Y_B^{(1)}(0), Y_B^{(2)}(0))$ and $\text{Cov}(Y_B^{(1)}(T), Y_B^{(2)}(T))$, each of which is non-positive, so we have that $\text{Cov}(Y_B^{(1)}(t^*), Y_B^{(2)}(t^*)) \leq 0$ for each interior interval boundary point t^* . The claim holds at the remaining points $t \in [0, T]$ by (5), since

$$\begin{aligned} & \text{Cov}(Y^{(1)}(t), Y^{(2)}(t)) \\ &= \frac{k}{2L} - t \left(\frac{k}{2L} + t - 2 \frac{k-1}{2L} \right) \text{Cov}(Y^{(1)}\left(\frac{k-1}{2L}\right), Y^{(2)}\left(\frac{k}{2L}\right)) \\ &+ \frac{(t - \frac{k-1}{2L})^2}{2^{-2L}} \text{Cov}(Y^{(1)}\left(\frac{k-1}{2L}\right), Y^{(2)}\left(\frac{k-1}{2L}\right)), \end{aligned} \quad (21)$$

for $t \in [\frac{k-1}{2L}, \frac{k}{2L}]$.

III. ANTICORRELATED CONTINUOUS-TIME MARKOV JUMP PROCESSES

We now produce variance reduced sample paths of continuous-time Markov jump processes by driving these processes with anticorrelated unit-rate Poisson processes. Consider the random time change representation of a Markov jump process,

$$X(t) = x_0 + \sum_{i=1}^I Y^i \left(\int_{s=0}^t \rho^i(s, X(s)) ds \right) \zeta^i,$$

where Y is a unit-rate Poisson process and ρ^i and ζ^i are the propensity function and jump vector of a reaction channel i . There are several ways to simulate such systems, for example Gillespie’s stochastic simulation algorithm (SSA) [3] for exact simulation or tau-leaping [4] for discrete-time approximate simulation. Another way to exactly simulate the system is to simulate the unit-rate Poisson processes that drive the Markov process and compute the realized jump times and states using the propensity functions and jump vectors. Taking this approach, if we replace the use of i.i.d. unit-rate Poisson processes with anticorrelated identically distributed unit-rate Poisson processes, as constructed in Section II, we produce exact realizations of the Markov jump process that are negatively correlated and hence produce variance-reduced mean estimates. Theorem 5 proves this result in the case of non-negative time-dependent propensity functions, $\rho^i(s, X(s)) = \rho^i(s)$.

Theorem 5: For each $i \in \{1, \dots, I\}$, let $Y^{i,(1)}(t)$ and $Y^{i,(2)}(t)$ be i.i.d. realizations of a unit-rate Poisson process Y . Let $Y_\alpha^{i,(1)}(t)$ and $Y_\alpha^{i,(2)}(t)$ be antithetically simulated realizations of Y using any of the above techniques such that $\text{Cov}(Y_\alpha^{i,(1)}(t), Y_\alpha^{i,(2)}(t)) \leq 0$ for every $t \in [0, T]$ and $Y_\alpha^{i_1, (j_1)}(t)$ and $Y_\alpha^{i_2, (j_2)}(t)$ are independent for each $i_1 \neq i_2$. Let $X^{(1)}(t)$ and $X^{(2)}(t)$ be i.i.d. realizations of a Markov jump process $X(t)$ such that

$$X^{(j)}(t) = x_0 + \sum_{i=1}^I Y^{i,(j)} \left(\int_{s=0}^t \rho^i(s) ds \right) \zeta^i, \quad (22)$$

for every $t \in [0, T^*]$ where T^* is defined such that $T = \max_i \int_{s=0}^{T^*} \rho^i(s) ds$. Let $X_\alpha^{(1)}(t)$ and $X_\alpha^{(2)}(t)$ be defined such that

$$X_\alpha^{(j)}(t) = x_0 + \sum_{i=1}^I Y_\alpha^{i,(j)} \left(\int_{s=0}^t \rho^i(s) ds \right) \zeta^i. \quad (23)$$

Finally, let $\Psi(t) := \frac{1}{2}(X^{(1)}(t) + X^{(2)}(t))$ and $\Psi_\alpha(t) := \frac{1}{2}(X_\alpha^{(1)}(t) + X_\alpha^{(2)}(t))$ be the unbiased 2-sample mean estimators of $X(t)$. Then $\text{tr Var}(\Psi_\alpha(t)) \leq \text{tr Var}(\Psi(t))$ for every $t \in [0, T^*]$. In particular, continuous time anticorrelated Markov jump process sampling reduces MSE of mean estimators.

Proof: It suffices to show that $\text{tr Cov}(X_\alpha^{(1)}(t), X_\alpha^{(2)}(t)) \leq 0$ for every $t \in [0, T^*]$. ■

Then

$$\begin{aligned} & \text{Cov}(X_\alpha^{(1)}(t), X_\alpha^{(2)}(t)) \\ &= \text{Cov} \left(\sum_{i_1=1}^I Y_\alpha^{i_1, (1)} \left(\int_{s=0}^t \rho^{i_1}(s) ds \right) \zeta^{i_1}, \right. \\ & \quad \left. \sum_{i_2=1}^I Y_\alpha^{i_2, (2)} \left(\int_{s=0}^t \rho^{i_2}(s) ds \right) \zeta^{i_2} \right) \quad (24) \\ &= \sum_{i=1}^I \text{Cov} \left(Y_\alpha^{i, (1)} \left(\int_{s=0}^t \rho^i(s) ds \right), \right. \\ & \quad \left. Y_\alpha^{i, (2)} \left(\int_{s=0}^t \rho^i(s) ds \right) \right) \zeta^i \zeta^{i\top}. \quad (25) \end{aligned}$$

The scalar covariance in (25) is negative by the Poisson process Lemmas above. Thus

$$\begin{aligned} & \text{tr Cov}(X_\alpha^{(1)}(t), X_\alpha^{(2)}(t)) \\ &= \sum_{i=1}^I \text{Cov} \left(Y_\alpha^{i, (1)}(\tau^i(t)), Y_\alpha^{i, (2)}(\tau^i(t)) \right) \|\zeta^i\|_2^2 \quad (26) \\ &\leq 0, \quad (27) \end{aligned}$$

for every $t \in [0, T^*]$, where $\tau^i(t) := \int_{s=0}^t \rho^i(s) ds$. ■

IV. RADIOACTIVE DECAY

While the requirement of Theorem 5 that the rate functions depend only on time can sometimes be restrictive, this condition is sufficient but hardly necessary. In practice, these variance reduction techniques can be beneficial even in the state-dependent or nonlinear cases. As an example, consider a simple model of radioactive decay with state-dependent rates, defined by

$$X(t) = x_0 - Y \left(\int_0^t \lambda X(s) ds \right), \quad (28)$$

for some scalar rate parameter $\lambda > 0$. A pair of anticorrelated sample paths of this process are shown in Fig. 9. The antithetic 4-sample mean estimator using each of the above techniques were constructed and their estimated autocovariances are shown in Fig. 10. Note the almost 75% variance reduction in some regimes.

REFERENCES

- [1] D. F. Anderson. An efficient finite difference method for parameter sensitivities of continuous time Markov chains. *SIAM Journal on Numerical Analysis*, 50:2237–2258, 2012.
- [2] C. Briat and M. Khammash. Computer control of gene expression: robust setpoint tracking of protein mean and variance using integral feedback. *IEEE Conference on Decision and Control*, 2012.
- [3] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.
- [4] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics*, 115(4):1716–1733, 2001.
- [5] T. G. Kurtz. Approximation of population processes. *CBMS-NSF Regional Conf. Ser. in Appl. Math*, 1981.
- [6] P. A. Maginnis. Variance reduction for Poisson and Markov jump processes. Master’s thesis, University of Illinois at Urbana-Champaign, 2011.

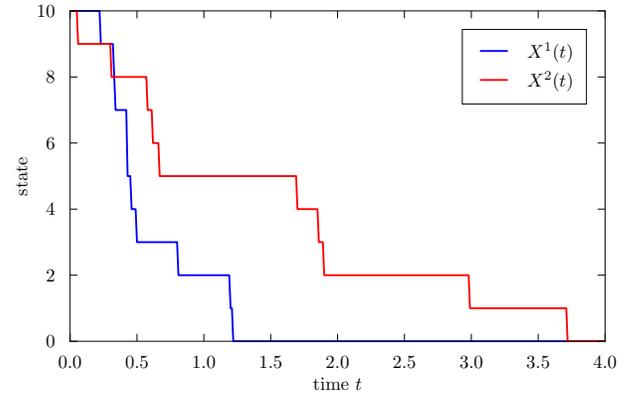


Fig. 9. A pair of sample paths of the radioactive decay process, driven by binomial midpoint antithetic Poisson processes.

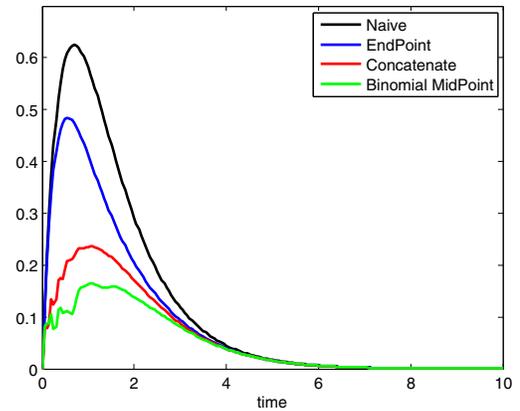


Fig. 10. Variance comparison 4-sample variance reduced mean estimators of a radioactive decay process using each type of antithetic sample path, with $K = 16/L = 4$.

- [7] P. A. Maginnis, M. West, and G. E. Dullerud. Application of variance reduction techniques for tau-leaping systems to particle filters. *IEEE Conference on Decision and Control*, 2012.
- [8] P. A. Maginnis, M. West, and G. E. Dullerud. Anticorrelated discrete-time stochastic simulation. *IEEE Conference on Decision and Control*, 2013.
- [9] M. Rathinam, P. Sheppard, and M. Khammash. Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks. *Journal of Chemical Physics*, 132(3):034103, 2010.
- [10] N. Riemer, M. West, R. A. Zaveri, and R. C. Easter. Simulating the evolution of soot mixing state with a particle-resolved aerosol model. *Journal of Geophysical Research*, 114, 2009.
- [11] C. P. Robert and G. Casella. *Monte Carlo statistical methods*. Springer, 2nd edition, 2004.