Proceedings of the 2008 Winter Simulation Conference
S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.

SIMULATING POINT PROCESSES BY INTENSITY PROJECTION

Kay Giesecke
Hossein Kakavand
Mohammad Mousavi

Department of Management Science and Engineering
Terman Center, Stanford University
Stanford, CA 94305, U.S.A.

ABSTRACT

Point processes with stochastic intensities are ubiquitous in many application areas, including finance, insurance, reliability and queuing. They can be simulated from standard Poisson arrivals by time-scaling with the cumulative intensity, whose path is typically generated with a discretization method. However, discretization introduces bias into the simulation results. This paper proposes a method for the exact simulation of point processes with stochastic intensities. The method leads to unbiased estimators. It is illustrated for a point process whose intensity follows an affine jump-diffusion process.

1 INTRODUCTION

Stochastic point processes are prominent in many application areas. In finance, they describe the arrival of economic events that are relevant to financial markets, such as corporate bankruptcies, mergers and acquisitions, or announcements of the Federal Reserve. In insurance, they describe claim arrivals. In reliability, they model equipment or software failures. In queuing applications, they record the arrivals or departures of customers. In health care, they represent incidences of infectious diseases. In seismology, they model earthquakes.

Monte Carlo simulation is an important computational tool to address point process applications. Many algorithms for generating point process sample paths exploit the specific structure of the intensity, which represents the conditional event arrival rate and governs the distribution of the point process. For example, if the intensity is deterministic between arrivals, then the thinning algorithm of (Lewis and Shedler 1979) applies, and arrivals can be generated from a dominating Poisson process by acceptance-rejection sampling. This approach leads to a relatively efficient simulation algorithm for a point process whose intensity depends at most on past event times and their marks, such as Poisson, birth and (Hawkes 1971) processes.

Typically, the thinning algorithm does not apply to a point process whose intensity follows a stochastic process between arrivals, since a dominating Poisson process may not exist. Yet such point processes are dominant in applications, in which arrival rates are often modulated by stochastic risk factors that follow complex dynamics on their own. A simulation approach that is applicable in this general case is based on the time change theorem of (Meyer 1971), which implies that under mild technical conditions a general point process with stochastic intensity can be transformed into a standard Poisson process by a change of time that is given by the point process compensator, or cumulative intensity. Thus, point process arrivals can be generated by re-scaling standard Poisson arrivals with the compensator. This approach requires the simulation of the continuous-time intensity process. If this process is approximated by a discrete-time process, then the simulation results suffer from discretization bias. This bias is undesirable for several reasons. First, since the size of the bias is unknown, it is hard to obtain valid confidence intervals. Second, a very fine time discretization may be required to reduce the bias to an acceptable level. Even more computational effort may be required to verify that the bias is sufficiently small.

In this article, we propose an *exact* simulation method that avoids the generation of intensity paths and leads to unbiased simulation results. The method is broadly applicable to point processes with stochastic intensities, and does not rely on the specific structure of the intensity at hand. It is based on the following result: we show that the arrival times of a point process N with intensity λ have the same joint distribution as the arrival times of a point process H whose intensity in the natural filtration of H is given by the optional projection of λ onto the natural filtration of N. The projected intensity is deterministic between arrivals, and therefore facilitates the sequential application of the classic thinning algorithm to generate event times by acceptance-rejection sampling. Our projection result guarantees that these event times have the correct joint distribution.

978-1-4244-2708-6/08/$25.00

We illustrate our exact simulation method for the family of affine point processes described in (Errais, Giesecke, and Goldberg 2006). The intensity of an affine point process follows an affine jump-diffusion process in the sense of (Duffie, Pan, and Singleton 2000). It jumps at event times, and undergoes diffusive fluctuations between events that are driven by a Brownian motion. Affine point processes are rich and versatile models for correlated event arrivals. The Poisson, birth and (Hawkes 1971) processes are special cases. Using results from point process filtering theory, we show how to explicitly calculate the projected intensity that the algorithm requires. A numerical case study illustrates the algorithm.

The remainder of this paper is organized as follows. Section 2 introduces the notation and recalls conventional approaches to point process simulation. Section 3 states and proves the projection result and describes the simulation algorithm it supports. Section 4 calculates the intensity projection for affine point processes and shows how to implement the simulation algorithm. Section 5 provides numerical results. Section 6 concludes.

2 POINT PROCESS

Consider a non-explosive counting process N with event stopping times (T_n), defined on a complete probability space $(\Omega, \mathscr{F}, P)$ with right-continuous and complete filtration $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ that represents the information flow. Suppose N has intensity process $\lambda = (\lambda_t)_{t\geq 0}$ relative to $\mathbb{F}$. This means that the process $N - A$ is a local martingale relative to $\mathbb{F}$, where $A = \int_0^{\cdot} \lambda_s ds$ is the $\mathbb{F}$-compensator to N. The intensity represents the conditional event arrival rate, in the sense that $\Delta\lambda_t$ is approximately equal to $P(N_{n+\Delta} - N_t = 1 \,|\, \mathscr{F}_t)$ for "small" Δ.

Let (ℓ_n) be a sequence of $\mathscr{F}_{T_n}$-measurable random variables with $\ell_0 = T_0 = 0$. The "mark" ℓ_n encodes additional information that is revealed at the event time T_n. For example, if the T_n model insurance claim arrival times, then the ℓ_n can describe the claim sizes. In portfolio credit risk applications, the T_n represent firm default times and the ℓ_n model the loss due to default. For concreteness, we suppose the ℓ_n are real-valued. Define the point process L by

$$L_t = \sum_{n=0}^{N_t} \ell_n = \sum_{n\geq 0} \ell_n 1_{\{T_n \leq t\}}. \tag{1}$$

We wish to simulate a trajectory of L, i.e. the sequence of pairs (T_n, ℓ_n). If the intensity λ is deterministic, then N is a non-homogeneous Poisson process and we can generate the arrival times by the inverse method from the inter-arrival time distribution, the order statistics property of the Poisson process, or the thinning algorithm of (Lewis and Shedler 1979). While fundamental, deterministic intensity models are often too simplistic. Many applications require state-dependent intensities. Here, λ follows an $\mathbb{F}$-adapted stochastic process. The value λ_t may depend on the path $\{L_s : 0 \leq s \leq t\}$, i.e. past arrival times and jumps $(T_n, \ell_n)_{n\leq N_t}$, and other sources of randomness recorded by the sigma-field $\mathscr{F}_t$. In the special case where λ is adapted to the filtration generated by L, i.e. if λ_t is a function only of past arrival times and jumps, then the inter-arrival intensity is deterministic, and the inter-arrival times may be generated sequentially by the inverse method from the conditional distribution of the inter-arrival time, or, preferably, thinning. Examples include birth and Hawkes processes (Ogata 1981).

In the general case, λ is modulated by additional random factors that follow stochastic processes on their own. In this case, λ is not adapted to the filtration generated by L anymore. Then, the inverse method is feasible only if the inverse of the conditional distribution function of the inter-arrival time can be evaluated, and the intensity values at the arrival times can be simulated. The required conditional distributions of the intensity and inter-arrival time are typically hard to evaluate except for particular families of inter-arrival intensity processes. (Giesecke and Kim 2007) consider such an example. While the thinning algorithm avoids the numerical inversion of the conditional inter-arrival time distribution, it applies only if the (random) inter-arrival intensity can be bounded above by some constant, and the values of the intensity at the candidate times can be simulated. However, in many cases of interest an upper bound for the inter-arrival intensity does not exist. For example, consider the case where λ is a function of a Feller diffusion process, or more generally, an affine jump-diffusion process (Duffie, Pan, and Singleton 2000).

There are few simulation methods that are applicable in the general stochastic intensity case without restrictions on the stochastic process followed by the intensity. One such method is based on the converse to the time change theorem of (Meyer 1971). Meyer's theorem implies that if the $\mathbb{F}$-compensator $A_t = \int_0^t \lambda_s ds$ increases to ∞ almost surely, then the counting process N can be transformed into a standard Poisson process by a change of time that is given by the inverse to A. Conversely, a standard Poisson process can be transformed into a point process with compensator A by a change of time that is given by A. This result suggests that we can generate the event times (T_n) of N from standard Poisson arrivals by time-scaling the Poisson times with the simulated compensator A.

The time-scaling method requires the simulation of the continuous-time intensity process λ, whose time integral $\int_0^t \lambda_s ds$ serves as the time transformation. Unless the intensity process has a very special structure, it must be approximated on a discrete-time grid. That is, the time interval is discretized and the intensity dynamics are simulated on this discrete-time grid. However, the approximation of continuous-time processes by discrete-time processes introduces bias into the simulation results.

3 SIMULATION BY PROJECTION

We propose a method for the exact simulation of the point process L that avoids the sampling of intensity values and leads to unbiased simulation results. Our method is as broadly applicable as the time-scaling approach, and does not require the intensity to follow a particular stochastic process. Instead of simulating L from the original $\mathbb{F}$-intensity model λ, we propose to *project* L onto a sub-filtration of $\mathbb{F}$, and then simulate the arrival times (T_n) from their intensity relative to the sub-filtration. The sub-filtration is chosen such that the structure of the sub-filtration intensity supports the application of the classical thinning algorithm.

We project L onto its own right-continuous and complete filtration $\mathbb{G} = (\mathscr{G}_t)_{t\geq 0}$ generated by the sigma-fields $\sigma(L_s : s \leq t)$. This is the smallest filtration that is compatible with L. The intensity of the counting process N relative to the filtration $\mathbb{G}$ is a $\mathbb{G}$-adapted process h such that H has compensator $\int_0^{\cdot} h_s ds$ relative to $\mathbb{G}$. That is, the process $N - \int_0^{\cdot} h_s ds$ is a $\mathbb{G}$-local martingale. The $\mathbb{G}$-intensity is given by the optional projection of the $\mathbb{F}$-intensity λ onto the sub-filtration $\mathbb{G}$. The optional projection is a $\mathbb{G}$-adapted process that is unique up to indistinguishability, see (Dellacherie and Meyer 1982, Numbers 43–44). It satisfies

$$h_t = E(\lambda_t \,|\, \mathscr{G}_t) \tag{2}$$

almost surely, for each t. If λ is $\mathbb{G}$-adapted as for Poisson, birth, Hawkes and time-inhomogeneous Markov point processes, then the optional projection $h = \lambda$. If λ is not a priori $\mathbb{G}$-adapted, then the optional projection is non-trivial. Due to the special structure of the sub-filtration $\mathbb{G}$, the projection evolves deterministically between events, and jumps at event times. Intuitively, the sources of randomness influencing λ beyond the arrival times T_n and jump sizes ℓ_n are "averaged out" by the projection. More precisely, h takes the form

$$h_t = \sum_{n\geq 0} h_n(t) 1_{\{N_t = n\}} \tag{3}$$

almost surely, where h_n is the $\mathscr{G}_{T_n}$-measurable function defined by

$$h_n(t) = \frac{E\big[\lambda_t 1_{\{N_t=n\}} \,|\, \mathscr{G}_{T_n}\big]}{P\big[N_t = n \,|\, \mathscr{G}_{T_n}\big]} \tag{4}$$

for $t \geq T_n$ and $n \geq 0$. There exist measurable functions $f_0 : \mathbb{R}_+ \to \mathbb{R}_+$ and $f_n : \mathbb{R}_+ \times (\mathbb{R}_+ \times \mathbb{R})^n \to \mathbb{R}_+$ such that

$$h_n(t) = \begin{cases} f_0(t) & n = 0 \\ f_n(t, T_1, \ell_1, \ldots, T_n, \ell_n) & n \geq 1 \end{cases} \tag{5}$$

The *hazard functions* f_n play a key role.

Proposition 3.1. *Let λ be an intensity of the counting process N in the filtration $\mathbb{F}$ such that*

$$E\left[\int_0^t \lambda_s ds\right] < \infty$$

for all t. Let h be the optional projection of λ onto the right-continous and complete filtration $\mathbb{G}$ generated by L. Let (f_n) be the sequence of functions prescribed by h through equation (5). Suppose H is a counting process with arrival times (S_n) starting at 0, and with intensity given by

$$\begin{cases} f_0(\cdot) & H = 0 \\ f_n(\cdot, S_1, \ell_1, \ldots, S_n, \ell_n) & H = n \geq 1 \end{cases}$$

relative to the right-continuous and complete filtration generated by $\sum_{n=0}^{H} \ell_n$. Then the arrival times of N and H have the same joint distribution:

$$(T_1, T_2, \ldots) \overset{\mathscr{L}}{=} (S_1, S_2, \ldots).$$

Proof. First observe that

$$1_{\{N_t > n\}} = \sum_{0<s\leq t} \Delta 1_{\{N_s > n\}} = \int_0^t 1_{\{N_{s-}=n\}} dN_s \tag{6}$$

where $\Delta V_t = V_t - V_{t-}$ is the jump of the process V at t and $V_{t-} = \lim_{s\uparrow t} V_s$. With the integrability hypothesis on λ, the local $\mathbb{F}$-martingale $N - \int_0^{\cdot} \lambda_s ds$ is a $\mathbb{F}$-martingale. Then, taking $\mathscr{F}_{T_n}$-conditional expectation on both sides of equation (6) gives

$$P\big[N_t > n \,|\, \mathscr{F}_{T_n}\big] = E\left[\int_{T_n}^t 1_{\{N_s=n\}} \lambda_s ds \,\middle|\, \mathscr{F}_{T_n}\right]$$

for $t \geq T_n$. Then, by iterated expectations and Fubini's theorem, we get

$$\begin{aligned} P\big[N_t > n \,|\, \mathscr{G}_{T_n}\big] &= \int_{T_n}^t E\big[\lambda_s 1_{\{N_s=n\}} \,|\, \mathscr{G}_{T_n}\big] ds \\ &= \int_{T_n}^t h_n(s) P\big[N_s = n \,|\, \mathscr{G}_{T_n}\big] ds, \end{aligned}$$

where the second equality follows from formula (4). By differentiating with respect to t we get

$$\begin{aligned} -\partial_t P\big[N_t = n \,|\, \mathscr{G}_{T_n}\big] &= \partial_t P\big[N_t > n \,|\, \mathscr{G}_{T_n}\big] \\ &= h_n(t) P\big[N_t = n \,|\, \mathscr{G}_{T_n}\big] \end{aligned}$$

which is easily solved to give

$$P\big[T_{n+1} > t \,|\, \mathscr{G}_{T_n}\big] = P\big[N_t = n \,|\, \mathscr{G}_{T_n}\big] = \exp\left(-\int_{T_n}^{t} h_n(s)ds\right). \tag{7}$$

Equation (7) demonstrates that the $\mathscr{G}_{T_n}$-conditional distribution of T_{n+1} is determined by the function $h_n = f_n(\cdot, T_1, \ell_1, \ldots, T_n, \ell_n)$ governing the $\mathbb{G}$-intensity h of N on $\{N = n\}$. An analogous argument can be used to show that the corresponding conditional distribution of the $(n+1)$th arrival time S_{n+1} of the counting process H takes the form (7) as well. More precisely, let $\mathbb{H} = (\mathscr{H}_t)_{t\geq 0}$ be the right-continuous and complete filtration generated by $\sum_{n=0}^{H} \ell_n$. By hypothesis, H has $\mathbb{H}$-intensity equal to $f_n(\cdot, S_1, \ell_1, \ldots, S_n, \ell_n)$ on $\{H = n\}$. Therefore, we get

$$P\big[S_{n+1} > t \,|\, \mathscr{H}_{S_n}\big] = P\big[H_t = n \,|\, \mathscr{H}_{S_n}\big] = \exp\left(-\int_{S_n}^{t} f_n(s, S_1, \ell_1, \ldots, S_n, \ell_n)ds\right).$$

Assuming that the first $n \geq 0$ arrival times of N and H have the same joint distribution, this observation implies that also the first $(n+1)$ arrival times of N and H have the same joint distribution. Since $N_0 = H_0 = 0$ by construction, T_1 and S_1 have the same distribution, and induction over n then shows that $(T_1, \ldots, T_n)$ and $(S_1, \ldots, S_n)$ have the same distribution. Kolmogorov's extension theorem concludes the proof. □

Proposition 3.1 justifies a simulation algorithm for L that is based on the projected intensity h. In particular, it shows that we can generate the arrival times T_{n+1} from the projected inter-arrival intensities h_n, which are described by the hazard functions f_n through formula (5). Given $\mathscr{G}_{T_n}$, the waiting time to next event T_{n+1} is equal in distribution to the first jump time of a non-homogeneous Poisson process started at T_n with intensity given by the $\mathscr{G}_{T_n}$-measurable function $h_n(t)$. Therefore, we can use the inverse method to simulate T_{n+1} from the corresponding conditional distribution function

$$P\big[T_{n+1} > t \,|\, \mathscr{G}_{T_n}\big] = \exp\left(-\int_{T_n}^{t} h_n(s)ds\right).$$

This method requires us to evaluate the inverse function to the conditional distribution function. Alternatively, we can apply the classical thinning algorithm of (Lewis and Shedler 1979), which requires us to evaluate $h_n(t)$ only at candidate arrival times generated from a dominating process.

Suppose the function $h_n(t)$ is decreasing. Then a Poisson process started at T_n with intensity $h_n(T_n)$ can serve as a dominating process. In this case, the generation of candidate times (V_k^n) for T_{n+1} is straightforward: $V_1^n = T_n + \mathscr{E}_1$, where $\mathscr{E}_1$ is an exponential random variate with parameter $h_n(T_n)$, $V_2^n = V_1^n + \mathscr{E}_2$, where $\mathscr{E}_2$ is an exponential random variate with parameter $h_n(T_n)$, independent of $\mathscr{E}_1$, and so on. A candidate time V_k^n is accepted as T_{n+1} with probability $h_n(V_k^n)/h_n(T_n)$. This basic algorithm can be improved by re-defining the dominating Poisson process after each rejection of a candidate time. That is, we can increase the acceptance probability by an adaptive choice of the bound on h_n. This is especially meaningful if h_n decays quickly.

Algorithm 3.2. *Assume the function $h_n(\omega)$ is decreasing. To generate the $(n+1)$th event time given the previous event times and marks,*

(i) *Initialize $k = 1$ and $V_0^n(\omega) = T_n(\omega)$.*
(ii) *Draw $\mathscr{E}_k \sim Exp(h_n(V_{k-1}^n(\omega))(\omega))$.*
(iii) *Set $V_k^n(\omega) = T_n(\omega) + \mathscr{E}_1(\omega) + \ldots + \mathscr{E}_k(\omega)$.*
(iv) *Draw $U \sim U(0,1)$. If*

$$U(\omega) \leq \frac{h_n(V_k^n(\omega))(\omega)}{h_n(V_{k-1}^n(\omega))(\omega)}$$

then set $T_{n+1}(\omega) = V_k^n(\omega)$ and stop. Else advance k by 1 and go to Step (ii).

For clarity in the exposition we have not included a termination condition that may be imposed by a fixed simulation horizon. Algorithm 3.2 must be modified to cover the case where h_n is not necessarily decreasing.

Algorithm 3.3. *To generate the $(n+1)$th event time given the previous event times and marks,*

(i) *Initialize $t = T_n(\omega)$.*
(ii) *Find functions $B_t^n(\omega)$ and $C_t^n(\omega)$ such that*

$$h_n(t+s)(\omega) \leq B_t^n(\omega), \quad 0 \leq s \leq C_t^n(\omega).$$

(iii) *Draw $\mathscr{E} \sim Exp(B_t^n(\omega))$. If*

- *$\mathscr{E}(\omega) > C_t^n(\omega)$ then set $t = t + C_t^n(\omega)$ and go to Step (ii).*
- *$\mathscr{E}(\omega) \leq C_t^n(\omega)$ then draw $U \sim U(0,1)$. If*

$$U(\omega) \leq \frac{h_n(t + \mathscr{E}(\omega))(\omega)}{B_t^n(\omega)}$$

then set $T_{n+1}(\omega) = t + \mathscr{E}(\omega)$ and stop. Else set $t = t + \mathscr{E}(\omega)$ and go to Step (ii).

The specific behavior of the function h_n at hand will determine the choice of the bound B_t^n and the interval length C_t^n for which it is valid. If, for example, h_n is increasing, then B_t^n can be taken to be the interval end point $h_n(t + C_t^n)$ for some C_t^n. The optimal value of C_t^n depends on the slope of h_n: intuitively, the steeper h_n the smaller C_t^n.

The algorithms are simple and easy to implement. The crucial ingredients are the functions h_n. Below we show how to calculate these functions for a broad family of point processes that have many applications.

4 SIMULATING AFFINE POINT PROCESSES

We illustrate our simulation approach for the case where L is a one-dimensional affine point process in the sense of (Errais, Giesecke, and Goldberg 2006).

4.1 Affine point process

Let X be an $\mathbb{F}$-adapted Markov state process that is a strong solution to the stochastic differential equation

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t + \delta\,dL_t \tag{8}$$

where $X_0 \in \mathbb{R}$, W is a standard Brownian motion relative to the filtration $\mathbb{F}$, $\mu(X)$ is the drift process, $\sigma(X)$ is the volatility process, $\delta \geq 0$ is a sensitivity parameter and $L = \sum_{n\geq 0}^{N} \ell_n$, see (1). The jump sizes ℓ_n are drawn independently of one another and independently of W from a discrete distribution ν on $\mathbb{R}_+$ that has no mass at zero.

We assume that $X \geq 0$ and that the counting process N has intensity $\lambda = X$. Since X jumps at the event times T_n, this choice implies that N and L are self-exciting point processes: an event increases the likelihood of further events. The parameter δ controls the sensitivity of the intensity to events. The jump sizes ℓ_n and the intensity λ are positively correlated: the bigger the jump the larger the response of the intensity, so the larger the increase of the likelihood of further events. Between events, the intensity diffuses according to the Brownian motion W.

The self-exciting feature and the correlation between jump sizes and event times are important in many applications. In portfolio credit risk, for example, we interpret X as a risk factor process, N as the default counting process and L as a process that records financial loss due to default. Defaults feed back on the risk factor that represents the economic state. Default rates and loss rates are positively correlated. Firms are exposed to a common diffusive risk factor W, whose movements generate correlated changes in firms' conditional default probabilities.

We further assume that the Markov state process X is an affine jump-diffusion in the sense of (Duffie, Pan, and Singleton 2000):

$$\mu(x) = K_0 + K_1 x, \quad \sigma(x)^2 = H_0 + H_1 x \tag{9}$$

for constant coefficients such that $X \geq 0$ and $E[\int_0^t X_s ds]$ is finite for all t. Special cases include the Poisson process ($H_0 = H_1 = \delta = 0$), the linear birth process ($K_0 = K_1 = H_0 = H_1 = 0$), and the Hawkes process ($H_0 = H_1 = 0$).

While the conditional characteristic function of (N,L) and related transforms can be calculated explicitly (Errais, Giesecke, and Goldberg 2006), in many applications we require quantities that are hard to determine in terms of the characteristic function. Moreover, the numerical inversion of the characteristic function can be challenging, especially if the parameters of the state process are not fixed in advance, and the inversion procedure must deal with a wide range of different parameter values. This situation regularly occurs in the context of calibration problems, where an optimization loop requires the distribution of (N,L) for many different parameter sets. Therefore, we are led to consider alternative methods to obtain the distribution of (N,L) and related quantities of interest.

The time-scaling approach can be used to simulate (N,L) and estimate the quantities of interest. This method requires us to generate sample paths of the affine jump-diffusion process $\lambda = X$. The Euler or some other higher-order discretization scheme can be used to approximate the path of X between events. While easy to implement, this approach will lead to biased simulation results unless the diffusion coefficient $\sigma = 0$, in which case the discretization is unnecessary. Our exact simulation method avoids the time discretization and therefore provides unbiased estimators even in the case where σ is non-trivial.

The exact approach to the simulation of the affine point process L can be combined with the exact method for the simulation of affine jump-diffusion processes with constant jump intensities developed in (Broadie and Kaya 2006). This combination would facilitate the exact simulation of affine jump-diffusion processes Z with stochastic jump intensities that are affine functions of Z. There are important applications of such a method in option pricing.

4.2 Intensity projection

The exact simulation method requires us to calculate the intensity h of N with respect to the filtration $\mathbb{G}$ generated by L. Consider the $\mathbb{G}$-conditional Laplace transform of the affine jump-diffusion X, given by

$$M_t(z) = E\left[\exp(-zX_t)\,|\,\mathscr{G}_t\right] = \int_0^\infty \exp(-zx)\pi_t(dx) \tag{10}$$

for $z \geq 0$, where $\pi_t(dx)$ is the conditional distribution of X_t given $\mathscr{G}_t$. Since X_0 is a constant, we have $M_0(z) = \exp(-zX_0)$. If $H_0 = H_1 = 0$ in (9), then the diffusion coefficient in the SDE (8) vanishes and X becomes $\mathbb{G}$-adapted, in which case $M_t(z) = \exp(-zX_t)$ for all $t \geq 0$. Whenever the diffusion coefficient is non-zero the filter $M(z)$ is non-trivial. From equation (2), the $\mathbb{G}$-intensity of N satisfies

$$h_t = -\partial_z M_t(z)|_{z=0}$$

almost surely, for each t.

The conditional transform $M_t(z)$ evolves deterministically between events and is updated at events. In our setting, the observation process L and the state process X have common jumps, and the point process filtering results in (Kliemann, Koch, and Marchetti 1990) imply that $M_t(z)$ is the solution to the Kushner-Stratonovich equation that describes the time evolution of the filter $E[f(t,X_t)\,|\,\mathscr{G}_t]$ for bounded and measurable functions f. See also (Ceci and Gerardi 2006), and for an affine setting (Frey, Prosdocimi, and Runggaldier 2007). The filter equation splits into an equation for the time interval $[T_n, T_{n+1})$ and an updating map for the event time T_{n+1}. This structure allows us to design an explicit recursive scheme to calculate the functions $h_n(t)$ that the simulation algorithm requires. To facilitate the calculations we assume that the jump size ℓ_n of L is measurable with respect to $\mathscr{G}_{T_n-}$. Intuitively, this means that the jump size at T_n is revealed "just before" T_n.

4.2.1 Auxiliary transform

To facilitate the filter calculation, let Y be a unique solution to the stochastic differential equation

$$Y_t = x + \int_s^t \mu(Y_u)\,du + \int_s^t \sigma(Y_u)\,dW_u \tag{11}$$

for any initial condition (s,x), where the coefficient functions μ and σ are given by equation (9). Note that Y_t and X_t agree in distribution if $\delta = 0$ and $(s,x) = (0,X_0)$, i.e. if X does not jump and starts at the same point as Y. Consider the transform

$$\varphi_t(s,x,v) = E\left[\exp\left(-\int_s^t Y_u du\right) e^{-vY_t}\,\middle|\, Y_s = x\right] \tag{12}$$

for $s \le t$ and x, v non-negative. Proposition 1 in (Duffie, Pan, and Singleton 2000) gives technical regularity conditions on the coefficients of the affine functions μ and σ such that

$$\varphi_t(s,x,v) = \exp\big(a(s,t,v) - b(s,t,v)x\big) \tag{13}$$

where the coefficient functions satisfy the ordinary differential equations

$$\partial_s b(s,t,v) = -K_1 b(s,t,v) + \frac{1}{2}H_1 b(s,t,v)^2 - 1 \tag{14}$$

$$\partial_s a(s,t,v) = K_0 b(s,t,v) - \frac{1}{2}H_0 b(s,t,v)^2 \tag{15}$$

with boundary conditions $b(t,t,v) = v$ and $a(t,t,v) = 0$. These ODEs can be solved explicitly for certain choices of the coefficient parameters. One such case is analyzed in Section 5 below.

4.2.2 Filtering between arrivals

Between arrivals the $\mathbb{G}$-Laplace transform of X is given by

$$M_t(z) = \frac{\rho_t^n(z)}{\rho_t^n(0)} \quad \text{on} \quad \{N_t = n\} \tag{16}$$

where, for $s \le t$ and $v \ge 0$, we have

$$\begin{aligned}\rho_t^n(v) &= \int_0^\infty \varphi_t(T_n,x,v)\,\pi_{T_n}(dx)\\ &= \exp\big(a(T_n,t,v)\big)\int_0^\infty \exp\big(-b(T_n,t,v)x\big)\,\pi_{T_n}(dx)\\ &= \exp(a(T_n,t,v))M_{T_n}(b(T_n,t,v)),\end{aligned} \tag{17}$$

where the second line follows from formula (13) and the third line is due to the definition of the conditional Laplace transform of X, and where b and a satisfy the ODEs (14) and (15) above. Thus, on the set $\{N_t = n\}$, we get

$$M_t(z) = \frac{\exp(a(T_n,t,z))M_{T_n}(b(T_n,t,z))}{\exp(a(T_n,t,0))M_{T_n}(b(T_n,t,0))}.$$

Next we show how to determine the filter at event times.

4.2.3 Filtering at arrivals

The filter $M(z)$ jumps at events. At the $(n+1)$th arrival time, the jump in $M(z)$ is given by

$$\begin{aligned}&M_{T_{n+1}}(z) - M_{T_{n+1}^-}(z)\\ &\quad= \frac{E\big[X_{T_{n+1}^-}\exp(-z(X_{T_{n+1}^-} + \delta\ell_{n+1}))\,|\,\mathscr{G}_{T_{n+1}^-}\big]}{E\big[X_{T_{n+1}^-}\,|\,\mathscr{G}_{T_{n+1}^-}\big]} - M_{T_{n+1}^-}(z)\end{aligned}$$

where $M_{t-}(z) = \lim_{s\uparrow t} M_s(z)$. Since the jump size variable ℓ_{n+1} is measurable with respect to $\mathscr{G}_{T_{n+1}^-}$, we can write

$$M_{T_{n+1}}(z) = \frac{E\big[X_{T_{n+1}^-}\exp(-zX_{T_{n+1}^-})\,|\,\mathscr{G}_{T_{n+1}^-}\big]J_{n+1}(z)}{E\big[X_{T_{n+1}^-}\,|\,\mathscr{G}_{T_{n+1}^-}\big]}$$

where

$$J_{n+1}(z) = \exp(-z\delta\ell_{n+1}), \quad z \ge 0.$$

Next we use the representation of the filter between arrivals developed in Section 4.2.2 above. With equation (16),

$$M_{T_{n+1}}(z) = J_{n+1}(z)\frac{\partial_z M_{T_{n+1}^-}(z)}{\partial_z M_{T_{n+1}^-}(z)|_{z=0}} = J_{n+1}(z)\frac{\partial_z \rho^n_{T_{n+1}^-}(z)}{\partial_z \rho^n_{T_{n+1}^-}(z)|_{z=0}}.$$

Finally, formula (17) allows us to express $M_{T_{n+1}}$ in terms of the Laplace transform M_{T_n} at the previous event time:

$$M_{T_{n+1}}(z) = J_{n+1}(z)\frac{\partial_v \exp(a(T_n,T_{n+1},v))M_{T_n}(b(T_n,T_{n+1},v))|_{v=z}}{\partial_v \exp(a(T_n,T_{n+1},v))M_{T_n}(b(T_n,T_{n+1},v))|_{v=0}}.$$

4.2.4 Recursive intensity calculation

We obtain a recursive algorithm to compute the functions $h_n(t)$ that govern the $\mathbb{G}$-intensity h of N via formula (3):

At $T_0 = 0$ we have $M_0(z) = \exp(-zX_0)$

For $t \geq T_n$:

$$h_n(t) = \frac{-\partial_v \exp(a(T_n,t,v))M_{T_n}(b(T_n,t,v))|_{v=0}}{\exp(a(T_n,t,0))M_{T_n}(b(T_n,t,0))}$$

At $t = T_{n+1}$:

$$h_{n+1}(T_{n+1}) = \delta\ell_{n+1} - \frac{\partial_v^2 \exp(a(T_n,T_{n+1},v))M_{T_n}(b(T_n,T_{n+1},v))|_{v=0}}{\partial_v \exp(a(T_n,T_{n+1},v))M_{T_n}(b(T_n,T_{n+1},v))|_{v=0}}$$

Recursion, at T_{n+1} compute

$$M_{T_{n+1}}(z) = J_{n+1}(z)\frac{\partial_v \exp(a(T_n,T_{n+1},v))M_{T_n}(b(T_n,T_{n+1},v))|_{v=z}}{\partial_v \exp(a(T_n,T_{n+1},v))M_{T_n}(b(T_n,T_{n+1},v))|_{v=0}}$$

4.3 Extensions

The basic affine jump-diffusion specification (8)–(9) for the state process X can be extended to include additional jump or diffusion terms that describe other stochastic risk factors that modulate the intensity. More generally, X can be a multi-dimensional affine jump-diffusion. Further, the coefficient functions (9) of X can be time-dependent. The intensity can be a time-dependent affine function of X. The results in (Duffie, Pan, and Singleton 2000) imply that the key transform (12) of the corresponding auxiliary process Y will retain its convenient exponentially affine form. Therefore, the filtering steps and the recursive algorithm for the calculation of the projected intensity remain valid with minor modifications.

5 NUMERICAL EXAMPLE

We implement the simulation procedure for a specific affine point process. We consider the special case where the coefficient functions (9) satisfy

$$(K_0, K_1, H_0, H_1) = (\kappa c, -\kappa, 0, \sigma^2) \tag{18}$$

for parameters $\kappa \geq 0$, $c > 0$ and $\sigma \geq 0$ such that $2\kappa c \geq \sigma^2$. In this case, the $\mathbb{F}$-intensity $\lambda = X$ of N follows an $\mathbb{F}$-Feller jump-diffusion process:

$$d\lambda_t = \kappa(c - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t + \delta dL_t. \tag{19}$$

The intensity jumps at the event times (T_n), with random jump sizes $(\delta\ell_n)$. Thus, the point processes N and L are self-exciting. The parameter δ controls the sensitivity of the arrival rate to events. After an event, λ reverts to the level c exponentially in mean at rate κ, and with diffusive fluctuations whose volatility is governed by σ. The auxiliary process Y introduced in equation (11) follows a classical $\mathbb{F}$-Feller diffusion.

In the special case $\sigma = 0$, the process L is a Hawkes process, while for $\kappa = \sigma = 0$ it is a linear birth process. In these two cases, our simulation method reduces to the sequential thinning method for Hawkes processes described in (Ogata 1981). For nontrivial σ, the exact Algorithm 3.2 is an alternative to the exact algorithm for the model (19) developed by (Giesecke and Kim 2007).

With the parametrization (18), the ODEs (14)–(15) take the convenient form

$$\partial_s b(s,t,v) = \kappa b(s,t,v) + \frac{1}{2}\sigma^2 b(s,t,v)^2 - 1$$
$$\partial_s a(s,t,v) = \kappa c b(s,t,v)$$

with boundary conditions $b(t,t,v) = v$ and $a(t,t,v) = 0$. We obtain the solutions

$$b(s,t,v) = \frac{v(\gamma+\kappa+(\gamma-\kappa)c(s,t)) + 2(c(s,t)-1)}{v\sigma^2(c(s,t)-1) + \gamma - \kappa + (\gamma+\kappa)c(s,t)}$$
$$a(s,t,v) = \frac{2\kappa c}{\sigma^2}\log\frac{2\gamma\exp((t-s)(\gamma+\kappa)/2)}{v\sigma^2(c(s,t)-1) + \gamma - \kappa + (\gamma+\kappa)c(s,t)}$$

for $\gamma = (\kappa^2 + 2\sigma^2)^{1/2}$ and $c(s,t) = \exp(\gamma(t-s))$. The projected intensity h can be computed explicitly using the recursive algorithm outlined in Section 4.2.4 above.

Figure 1 shows sample paths of the projected intensity h and the point process L. The jumps in h occur at the event times. Between events h decreases deterministically. Therefore, Algorithm 3.2 is applicable. The parameters of

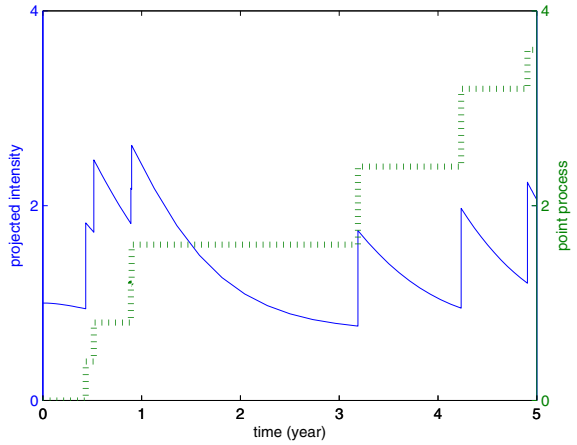

Figure 1: Sample paths of the projected intensity h (left scale, solid line) and the point process L (right scale, dashed line) generated by the exact Algorithm 3.2. The jumps in h occur at the event times. Between events h decays deterministically. The reversion rate $\kappa = 1$, the reversion level $c = 1$, the diffusive volatility $\sigma = 1$, the feedback sensitivity $\delta = 1$ and the initial value $X_0 = 1$. The jump size variables ℓ_n are drawn from a uniform distribution ν over $\{0.4, 0.8\}$.

the $\mathbb{F}$-intensity λ are given by

$$(\kappa, c, \sigma, \delta, X_0) = (1, 1, 1, 1, 1)$$

and the jump size variables ℓ_n are drawn from a uniform distribution ν over $\{0.4, 0.8\}$.

The numerical experiments were performed on a laptop PC with an Intel 1.6 GHz processor and 1 GB of RAM, running Windows XP Home. The codes were written in MATLAB Version 7.2.0 (R2006a).

6 CONCLUSION

This paper proposes a new method for the exact simulation of point processes with stochastic intensities. The method avoids the generation of intensity paths and leads to unbiased estimators. It is based on the observation that the joint distribution of the point process arrival times is completely determined by the optional projection of the intensity onto the filtration generated by the point process itself. The optional projection is the intensity of the point process in its own filtration. Between events, it evolves deterministically, and therefore supports the application of the classical thinning algorithm to efficiently generate event times by acceptance-rejection sampling. We demonstrate the method for a point process whose intensity follows an affine jump-diffusion.

ACKNOWLEDGMENTS

This research was supported by a grant from Mizuho-DL Financial Technology, for which we are very grateful. We would like to thank Hideyuki Takada and Stefan Weber for helpful discussions.

REFERENCES

Broadie, M., and O. Kaya. 2006. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* 54 (2): 217–231.

Ceci, C., and A. Gerardi. 2006. A model for high-frequency data under partial information: a filtering approach. *International Journal of Theoretical and Applied Finance* 9 (4): 555–576.

Dellacherie, C., and P.-A. Meyer. 1982. *Probabilities and potential*. Amsterdam: North Holland.

Duffie, D., J. Pan, and K. Singleton. 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68:1343–1376.

Errais, E., K. Giesecke, and L. Goldberg. 2006. Pricing credit from the top down with affine point processes. Working Paper, Stanford University.

Frey, R., C. Prosdocimi, and W. Runggaldier. 2007. Affine credit risk models under incomplete information. In *Stochastic Processes and Applications to Mathematical Finance - Proceedings of the 6th Ritsumeikan International Symposium*, ed. J. Akahori, S. Ogawa, and S. Watanabe, 97–113. World Scientific.

Giesecke, K., and B. Kim. 2007. Estimating tranche spreads by loss process simulation. In *Proceedings of the 2007 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 967–975. IEEE Press.

Hawkes, A. G. 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58 (1): 83–90.

Kliemann, W., G. Koch, and F. Marchetti. 1990. On the unnormalized solution of the filtering problem with counting process observations. *IEEE Transactions on Information Theory* 36 (6): 1415–1425.

Lewis, P., and G. Shedler. 1979. Simulation of nonhomogeneous poisson processes by thinning. *Naval Logistics Quarterly* 26:403–413.

Meyer, P.-A. 1971. Démonstration simplifée d'un théorème de Knight. In *Séminaire de Probabilités V, Lecture Note in Mathematics 191*, 191–195. Springer-Verlag Berlin.

Ogata, Y. 1981. On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory* 27:23–31.

AUTHOR BIOGRAPHIES

KAY GIESECKE is an assistant professor in the Department of Management Science and Engineering at Stanford University. He has previously held a position in the School of Operations Research and Information Engineering at Cornell University. His research is in financial engineering, particularly credit risk modeling. His email is giesecke@stanford.edu and his web page is <www.stanford.edu/∼giesecke>.

HOSSEIN KAKAVAND is a postdoctoral fellow in the Department of Management Science and Engineering at Stanford University. His research interests include credit risk modeling, financial risk management, and computation theory. His email is hossein@stanford.edu and web page is <www.stanford.edu/∼hossein>.

MOHAMMAD MOUSAVI is a Ph.D. candidate in the Department of Management Science and Engineering at Stanford University. His research interests include financial engineering and stochastic modeling. His email is mousavi@stanford.edu and web page is <www.stanford.edu/∼mousavi>.