Proceedings of the 2009 Winter Simulation Conference
M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, eds.

RARE EVENT SIMULATION FOR A GENERALIZED HAWKES PROCESS

Xiao-Wei Zhang
Peter W. Glynn
Kay Giesecke

Department of Management Science
and Engineering
Stanford University
Stanford, C.A. 94305, U.S.A.

Jose Blanchet

Department of Industrial Engineering
and Operations Research
Columbia University
New York, N.Y. 10027, U.S.A.

ABSTRACT

In this paper we study rare event simulation for the tail probability of an affine point process $(J_t)_{t\geq 0}$ that generalizes the Hawkes process. By constructing a suitable exponential martingale, we are able to construct an importance sampling algorithm that is logarithmically efficient in the Gartner-Ellis asymptotic regime.

1 INTRODUCTION

Affine point processes model have been used in the credit risk literature to capture the "clustering" or "self-exciting" feature of the credit defaults that have been observed in the financial industry. In such models, the number of defaults is modeled as a point process with an intensity driven by market-wide risk factors that follow an affine jump diffusion. The default counting process itself is a risk factor as well so that the timing of past defaults influences the future evolution of defaults; see Errais, Giesecke, and Goldberg (2009) for more discussion on the pricing and modeling of credit derivatives using affine point processes.

Consider the affine point process satisfying the following stochastic differential equation (SDE)

$$\mathrm{d}\lambda_t = \kappa(\mu - \lambda_t)\,\mathrm{d}t + \sigma\sqrt{\lambda_t}\,\mathrm{d}B_t + \delta\,\mathrm{d}J_t, \tag{1}$$

given $\lambda_0 > 0$ and $J_0 = 0$, where $(B_t : t \geq 0)$ is a standard Brownian motion and $(J_t : t \geq 0)$ is a counting process with intensity $(\lambda_t : t \geq 0)$, i.e. $\Lambda_t \triangleq \int_0^t \lambda_s\,\mathrm{d}s$ is the compensator of J_t or equivalently is such that $J_t - \Lambda_t$ is a local martingale. One may view J_t as the cumulative number of defaults by time t and λ_t is the associated arrival intensity of defaults. Moreover, at a default event, the intensity jumps by an amount δ.

The special case where $\sigma = 0$ (so that there is no Brownian noise in (1)) is known as the Hawkes process. We therefore refer to the model (1) as a generalized Hawkes process. Given that λ_t is an intensity, we require that $(\lambda_t : t \geq 0)$ be a non-negative process in order that $(J_t : t \geq 0)$ be well-defined. This imposes the condition that $2\kappa\mu \geq \sigma^2$ and consequently this condition will be in force throughout the remainder of this paper.

Suppose that we are interested in the probability distribution of the number of defaults by time t. In particular, suppose we want to compute $\mathrm{P}(J_t > x)$. It is well known that the Fourier transform of J_t has an exponential affine form which can be identified by solving (generalized) Riccati ordinary differential equations (ODE's); see Duffie, Pan, and Singleton (2000) and Errais, Giesecke, and Goldberg (2009). Since J_t is integer-valued, $\mathrm{P}(J_t = n)$ can be characterized directly in terms of ODEs. However, the ODEs take an inconvenient recursive structure, making it more and more difficult to get the probabilities $\mathrm{P}(J_t = n)$ for increasing n. In this paper, we focus on the use of Monte Carlo simulation, which can be potentially generalized to the multidimensional case without being affected by the "curse of dimensionality" associated with Fourier transform methods.

The main difficulty in computing the tail probability via crude Monte Carlo (CMC) is that the number of trials n required to estimate α, the probability of interest, to a given relative precision scales in rough proportion to α^{-1}. As a consequence, CMC is highly inefficient for estimating small α. Importance sampling is a technique that is widely used to reduce the variance of such estimators (thereby reducing the computational cost); see, for example, Bucklew (2004).

Suppose $(A_\gamma : \gamma \in \Gamma)$ is a family of rare events, i.e. $\inf_{\gamma\in\Gamma} \mathrm{P}(A_\gamma) = 0$. Let $(Z(\gamma) : \gamma \in \Gamma)$ be a family of unbiased estimators for the $\mathrm{P}(A_\gamma)$'s. Such a family of estimators is said to be *logarithmically efficient* if

$$\inf_{\gamma\in\Gamma} \frac{|\log(\mathrm{E}Z(\gamma)^2)|}{|\log(\alpha(\gamma)^2)|} \geq 1;$$

see Asmussen and Glynn (2007) for a detailed discussion.

In this paper, we study the rare event regime $\{J_t > xt\}$ as $t \to \infty$ for suitably large x. The structure of this paper is as follows. Section 2 introduces a family of change-of-measure martingales from which we will choose our importance sampling distribution. Section 3 calculates the logarithmic asymptotics of $\mathrm{P}(J_t > xt)$, specifying its exponential decay rate. Section 4 proposes an importance sampling algorithm and shows it is asymptotically optimal, with numerical results provided in Section 5. The proofs of the propositions are provided in Appendix.

2 A CLASS OF EXPONENTIAL MARTINGALES

Note that J_t is an additive functional of λ_t. This suggests that an appropriate importance distribution will be induced by an exponential martingale of the form

$$M_t(\theta) = \exp(h(\theta,\lambda_t) + \theta J_t - \psi(\theta)t - h(\theta,\lambda_0))$$

for $\theta > 0$ where $\psi(\theta)$ is deterministic and to be specified. Because λ_t is an affine process, it turns out that $h(\theta,\lambda_t)$ can be taken to be linear in λ_t, so that $h(\theta,\lambda_t) = a(\theta)\lambda_t$ for some suitable function $a(\theta)$. We now proceed to compute $\psi(\theta)$ and $a(\theta)$.

Itô's formula (see Chapter 2 of Protter 2005) establishes that

$$M_t := M_t(\theta) = 1 + \int_0^t M_{s-}\,\mathrm{d}Y_s + \frac{1}{2}\int_0^t M_{s-}\,\mathrm{d}[Y,Y]_s^c + \sum_{0<s\leq t}(M_s - M_{s-} - M_{s-}\cdot\Delta Y_s), \tag{2}$$

where $Y_t = a(\theta)\lambda_t + \theta J_t - \psi(\theta)t - a(\theta)\lambda_0$, $[Y,Y]^c$ is the path-by-path continuous part of the quadratic variation process $[Y,Y]$, $Y_{0-} = Y_0 = 0$ and $M_{0-} = M_0 = 1$. Note that

$$Y_t = (a(\theta)\kappa\mu - \psi(\theta))t + a(\theta)\sigma\int_0^t \sqrt{\lambda_s}\,\mathrm{d}B_s - a(\theta)\kappa\Lambda_t + (a(\theta)\delta+\theta)J_t, \tag{3}$$

so

$$[Y,Y]_t = a(\theta)^2\sigma^2\int_0^t \lambda_s\,ds + (a(\theta)\delta+\theta)^2 J_t.$$

It follows that $[Y,Y]_t^c = a(\theta)^2\sigma^2\Lambda_t$ and thus

$$\int_0^t M_{s-}\,\mathrm{d}[Y,Y]_s^c = a(\theta)^2\sigma^2\int_0^t M_{s-}\,\mathrm{d}\Lambda_s. \tag{4}$$

Let

$$A_t = (a(\theta)\kappa\mu - \psi(\theta))t + a(\theta)\sigma\int_0^t \sqrt{\lambda_s}\,\mathrm{d}B_s - a(\theta)\kappa\Lambda_t$$

be the continuous part of Y_t. Then,

$$\begin{aligned}\sum_{0<s\leq t}(M_s - M_{s-}) &= \sum_{0<s\leq t} e^{A_{s-}}(e^{(a(\theta)\delta+\theta)(J_{s-}+1)} - e^{(a(\theta)\delta+\theta)J_{s-}})\cdot\Delta J_s \\ &= (e^{(a(\theta)\delta+\theta)} - 1)\sum_{0<s\leq t} e^{A_{s-}+(a(\theta)\delta+\theta)J_{s-}}\cdot\Delta J_s = (e^{(a(\theta)\delta+\theta)} - 1)\int_0^t M_{s-}\,\mathrm{d}J_s.\end{aligned} \tag{5}$$

Moreover, we have

$$\sum_{0<s\le t} M_{s-}\cdot \Delta Y_s = (a(\theta)\delta+\theta)\sum_{0<s\le t} M_{s-}\cdot \Delta J_s = (a(\theta)\delta+\theta)\int_0^t M_{s-}\,\mathrm{d}J_s. \tag{6}$$

If we plug (3) - (6) into (2), we find that

$$\begin{aligned}\mathrm{d}M_t =&(a(\theta)\kappa\mu-\psi(\theta))M_{t-}\,\mathrm{d}t + a(\theta)\sigma\sqrt{\lambda_t}M_{t-}\,\mathrm{d}B_t + \left(e^{a(\theta)\delta+\theta}-1\right)M_{t-}(\mathrm{d}J_t-\mathrm{d}\Lambda_t)\\ &+\left(\frac{a(\theta)^2\sigma^2}{2}-a(\theta)\kappa+e^{a(\theta)\delta+\theta}-1\right)M_{t-}\lambda_t\,\mathrm{d}t.\end{aligned}$$

Since $(B_t : t\ge 0)$ and $(J_t-\Lambda_t : t\ge 0)$ are local martingales, it is evident that M_t will be a local martingale if we choose

$$a(\theta)\kappa\mu-\psi(\theta)=0 \tag{7}$$

and

$$\frac{a(\theta)^2\sigma^2}{2}-a(\theta)\kappa+e^{a(\theta)\delta+\theta}-1=0. \tag{8}$$

Observe that (7) asserts that $\psi(\theta)$ is computable from $a(\theta)$. Hence, the question of whether a suitable exponential martingale exists for a given θ has been reduced to the issue of whether (8) possesses a solution $a(\theta)$; we discuss this equation in Section 3.

Assuming the existence of $a(\theta)$ in (8), $M_t(\theta)$ is a local martingale. When $\{M_t : 0\le t\le T\}$ is actually a martingale, we may therefore define an equivalent probability measure $Q(\cdot)$ by $\frac{dQ}{dP}\big|_{\mathscr{F}_t}=M_t$. In order to identify the dynamics of (λ, J) under Q, observe that

$$Y_t=(\theta+a(\theta)\delta)J_t+\left(1-e^{\theta+a(\theta)\delta}\right)\int_0^t\lambda_s\,\mathrm{d}s-\frac{a(\theta)^2\sigma^2}{2}\int_0^t\lambda_s\,\mathrm{d}s+\int_0^t a(\theta)\sigma\sqrt{\lambda_s}\,\mathrm{d}B_s,$$

because of (7) and (8). It follows, letting

$$M_t^{(1)}\triangleq\exp\left((\theta+a(\theta)\delta)J_t+\left(1-e^{\theta+a(\theta)\delta}\right)\int_0^t\lambda_s\,\mathrm{d}s\right)$$

and

$$M_t^{(2)}\triangleq\exp\left(-\frac{a(\theta)^2\sigma^2}{2}\int_0^t\lambda_s\,\mathrm{d}s+\int_0^t a(\theta)\sigma\sqrt{\lambda_s}\,\mathrm{d}B_s\right),$$

that $M_t=M_t^{(1)}M_t^{(2)}$. By Girsanov's theorem (see Chapter 1 of Oksendal and Sulem 2007), M_t represents two changes-of-measure corresponding to the two sources of randomness: $M_t^{(1)}$ changes the intensity of the counting process while $M_t^{(2)}$ changes the drift of the Brownian motion for Itô processes. In particular, the dynamics of λ under Q is governed by the following SDE

$$\mathrm{d}\lambda_t=\kappa(\theta)(\mu(\theta)-\lambda_t)\,\mathrm{d}t+\sigma\sqrt{\lambda_t}\,\mathrm{d}B_t^Q+\delta\,\mathrm{d}J_t^Q, \tag{9}$$

where $\kappa(\theta)=\kappa-a(\theta)^2\sigma$, $\mu(\theta)=\kappa\mu/\kappa(\theta)$, $(B_t^Q : t\ge 0)$ is a standard Brownian motion under Q and J_t^Q follows a counting process with intensity $\lambda_t\exp(a(\theta)\delta+\theta)$.

3 LOGARITHMIC ASYMPTOTICS

In this section, we calculate the logarithmic asymptotics of $\mathrm{P}(J_t > xt)$ as well as the correct "exponential twisting" parameter θ^* which leads to an logarithmically efficient algorithm. We begin this section by computing the equilibrium mean of $(\lambda_t : t\ge 0)$.

Observe that if we formally take expectations in (1), we find that

$$\mathrm{d}\mathrm{E}\lambda_t = (\kappa\mu - (\kappa-\delta)\mathrm{E}\lambda_t)\,\mathrm{d}t,$$

thereby yielding a (deterministic) differential equation for $\mathrm{E}\lambda_t$. If $\kappa < \delta$, this differential equation implies that $\mathrm{E}\lambda_t$ increases exponentially in t, so that $\mathrm{P}(J_t > xt)$ is not a rare event. On the other hand, if $\kappa > \delta$, then $\mathrm{E}\lambda_t$ should converge exponentially rapidly to $\kappa\mu/(\kappa-\delta)$; this should correspond to the setting where $(\lambda_t : t \geq 0)$ is a recurrent Markov process so that events like $\{J_t > xt\}$ are rare. Not surprisingly, we can the establish the following result.

Proposition 1 *Suppose $\kappa > \delta$. Then, there exists $\widetilde{\theta} > 0$ and $a_1 > 0$ such that (8) has a positive root $a(\theta) \in [0, a_1)$ for $\theta \in [0, \widetilde{\theta})$.*

The proof can be found in the Appendix of this paper. The positivity restriction of $a(\theta)$ arises as a consequence of (7), since $\psi(\theta)$ must be positive in order that M_t be a martingale for $\theta > 0$.

When $\theta \in (0, \widetilde{\theta})$ and $a(\theta) \in (0, a_1)$, the process $(\lambda_t : t \geq 0)$ is geometrically ergodic under the change-of-measure Q; see Zhang, Glynn, Giesecke, and Blanchet (2009). This suggests that

$$\mathrm{E}\exp(\theta J_t - \psi(\theta)t) = \mathrm{E}^Q \exp(a(\theta)(\lambda_0 - \lambda_t)) \rightarrow \mathrm{E}^Q \exp(a(\theta)\lambda_\infty) \tag{10}$$

as $t \rightarrow \infty$, where $\mathrm{E}^Q(\cdot)$ is the expectation operator associated with Q, and λ_∞ has the equilibrium distribution of $(\lambda_t : t \geq 0)$; see Zhang, Glynn, Giesecke, and Blanchet (2009) for a complete argument.

Given our above formal calculation, $\mathrm{P}(J_t > xt)$ is a family of rare events as $t \rightarrow \infty$ for $x > \kappa\mu/(\kappa-\delta)$. Given the limit relationship (10),

$$\lim_{t\rightarrow\infty} \frac{1}{t} \log \mathrm{E}\exp(\theta J_t) = \psi(\theta). \tag{11}$$

It follows from the Gartner-Ellis theorem (see Dembo and Zeitouni 1998 or Bucklew 2004) that if there exists θ^* such that $\psi'(\theta^*) = x$, then

$$\lim_{t\rightarrow\infty} \frac{1}{t} \log \mathrm{P}(J_t > xt) = -I(x), \tag{12}$$

where $I(x) = \theta^* x - \psi(\theta^*)$; existence of such a θ^* is established in Proposition 2 with proof provided in the Appendix.

Proposition 2 *For each $x > \kappa\mu/(\kappa-\delta)$, θ^* is given by*

$$\theta^* = \theta^*(x) = -\delta a \log\left(1 + \kappa a - \frac{\sigma^2}{2} a^2\right)$$

where a is the smaller root of the quadratic equation

$$\left(\frac{\delta\kappa\mu}{\kappa-\delta} + 1\right)\frac{\sigma^2}{2} a^2 - \left(\frac{\kappa\mu(\sigma^2 + \kappa\delta)}{\kappa-\delta} + \kappa\right) a + \frac{(\kappa-\delta)x}{\kappa\mu} - 1 = 0.$$

Given the limit (12), this suggests the approximation

$$\mathrm{P}(J_t > xt) \approx \exp(-I(x)t) \tag{13}$$

when t is large. In practice, the approximation (13) can be quite poor for moderate values of t. It follows that use of simulation to compute $\mathrm{P}(J_t > xt)$ is of significant practical value in many settings.

4 THE IMPORTANCE SAMPLING ALGORITHM

Given the key role that θ^* plays in the asymptotic limit (13), it is natural to therefore consider an importance sampling algorithm in which the importance distribution is the change-of-measure Q associated with the exponential martingale associated with θ^*. For this choice of importance distribution, the associated estimator is

$$Z(t) \triangleq M_t(\theta^*)^{-1}\mathbb{I}(J_t^Q > xt).$$

The corresponding importance sampling algorithm is then given by:

i.) Compute the root θ^* of $\psi'(\theta^*) = x$

ii.) Simulate λ_t and J_t under the dynamics of the importance distribution Q associated with θ^* (see Giesecke and Kim 2007)

iii.) Compute $Z(t)$ from the simulated path

iv.) Replicate steps i.) – iii.) N iid times, thereby producing $Z_1(t),\dots,Z_N(t)$

v.) Calculate

$$\bar{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i(t)$$

as the estimate of $\mathrm{P}(J_t > xt)$.

Note that

$$\begin{aligned}
\mathrm{E}^Q Z^2(t) &= \mathrm{E}^Q \exp[2(\psi(\theta^*)t - \theta^* J_t - a(\theta^*)(\lambda_t - \lambda_0))]\mathbb{I}(J_t > xt) \\
&\le \mathrm{E}^Q \exp[2(\psi(\theta^*)t - \theta^* xt - a(\theta^*)(\lambda_t - \lambda_0))] \\
&= \exp(-2I(x)t)\mathrm{E}^Q \exp(-2a(\theta^*)(\lambda_t - \lambda_0))
\end{aligned}$$

The geometric ergodicity of $(\lambda_t : t \ge 0)$ under Q then suggests that

$$\overline{\lim_{t\to\infty}} \exp(2I(x)t)\mathrm{E}^Q Z^2(t) < \infty.$$

This guarantees that the family of estimators $(Z(t) : t \ge 0)$ is logarithmically efficient for computing the $\mathrm{P}(J_t > xt)$'s; see Zhang, Glynn, Giesecke, and Blanchet (2009) for details.

5 NUMERICAL RESULTS

The simulation experiments were performed on a desktop PC with an Intel Core 2 Quad 2.40 GHz processor and 2GB of RAM, running Windows XP Professional. The codes were written in C++. The compiler used was Microsoft Visual Studio 2008. The numerical results are shown in Table 1 and Figure 1.

A APPENDIX

A.1 Proof of Proposition 1

Rewrite (8) as

$$e^{\theta} = e^{-\delta a}\left(1 + \kappa a - \frac{\sigma^2}{2}a^2\right) \triangleq f(a).$$

We have

$$f'(a) = e^{-\delta a}\left(\kappa - \delta - (\sigma^2 + \kappa\delta)a + \frac{\delta\sigma^2}{2}a^2\right) \triangleq e^{-\delta a}g(a).$$

Table 1: Simulation results for $P(J_t > xt)$. Parameter values are $\kappa = 5$, $\mu = 0.7$, $\sigma = 0.2$, $\delta = 1$, $\lambda_0 = 0.7$ and $x = 1.2$. "Log ratio" means $\log \mathrm{E}^Q Z^2(t)/\log(\mathrm{P}(J_t > xt)^2)$. Number of simulation trials is 500.

t	$Z(t)$	$\mathrm{E}^Q Z(t)^2$	log ratio	95% CI of log ratio
100	3.83E-03	6.57E-05	0.865177	[0.856314,0.87404]
200	1.30E-04	8.39E-08	0.910286	[0.90366,0.916911]
300	3.57E-06	8.70E-11	0.92335	[0.917782,0.928918]
400	1.04E-07	8.89E-14	0.934571	[0.92959,0.939552]
500	3.31E-09	1.01E-16	0.943133	[0.938622,0.947645]
700	4.47E-12	1.70E-22	0.959022	[0.955666,0.962378]
1000	1.39E-16	2.66E-31	0.964063	[0.960877,0.967248]
1400	2.93E-22	1.12E-42	0.974118	[0.971661,0.976574]
2000	6.07E-31	5.69E-60	0.980341	[0.978514,0.98169]
3500	1.48E-52	5.32E-103	0.986644	[0.985196,0.988091]
5000	6.58E-74	7.43E-146	0.991564	[0.990698,0.99243]
7500	4.34E-110	7.39E-218	0.992712	[0.990572,0.994852]
10000	6.08E-146	1.30E-289	0.99468	[0.993156,0.996204]

The discriminant of $g(a)$

$$\Delta_g = \sigma^4 + 2\sigma^2\delta^2 + \kappa^2\delta^2 > (\sigma^2+\delta^2)^2 \geq 0$$

since $\kappa > \delta$. Hence $g(a)$, as well as $f'(a)$, has two distinct positive zeros, say $a_1 < a_2$. It follows that

$$\sup_{a>0} f(a) = f(a_1) > f(0) = 1, \quad \inf_{a>0} f(a) = f(a_2) < 0.$$

Moreover, $f(a)$ is strictly increasing on $(0,a_1)$ and strictly decreasing on (a_1,a_2). Let $\widetilde{\theta} = \log(f(a_1))$. Then $f(a) = e^\theta$ has two distinct roots for $\theta < \widetilde{\theta}$, one single root for $\theta = \widetilde{\theta}$ and no root for $\theta > \widetilde{\theta}$. When $\theta < \widetilde{\theta}$, we choose $a(\theta) < a_1$ to be the smaller root.

A.2 Proof of Proposition 2

Put $y = x/(\kappa\mu) > 0$, then we only need to solve (8) and $a'(\theta^*) = y$. Differentiating w.r.t. θ on both sides of (8), we have

$$a(\theta^*)\sigma^2 y - \kappa y + (\delta y + 1)e^{a(\theta^*)\delta + \theta^*} = 0 \tag{14}$$

Combing (14) with (8) yields

$$(\delta y+1)\frac{\sigma^2}{2}a(\theta^*)^2 - ((\sigma^2+\kappa\delta)y+\kappa)a(\theta^*) + (\kappa-\delta)y - 1 = 0 \tag{15}$$

which is a quadratic in $a(\theta^*)$. Call the LHS of (15) $l(a)$, i.e.

$$l(a) = (\delta y+1)\frac{\sigma^2}{2}a^2 - ((\sigma^2+\kappa\delta)y+\kappa)a + (\kappa-\delta)y - 1.$$

Then, (15) has solutions for $a(\theta^*)$ if and only if the discriminant of $l(a)$

$$\Delta_l \equiv ((\sigma^2+\kappa\delta)y+\kappa)^2 - 2\sigma^2(\delta y+1)((\kappa-\delta)y-1) \geq 0.$$

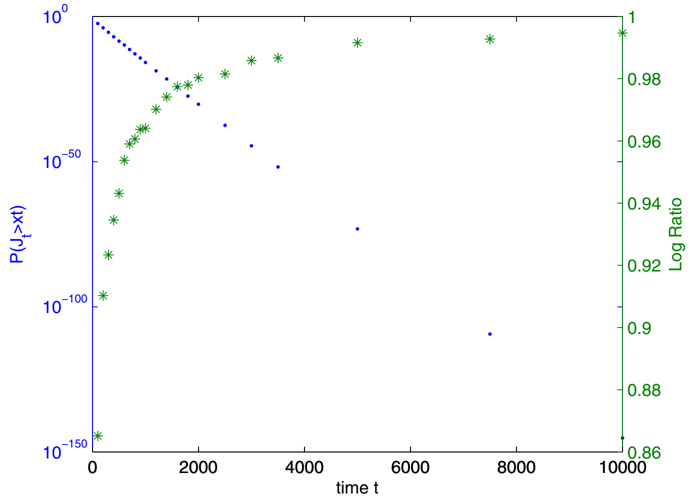

Figure 1: Plot of log ratio as function of time t

Indeed, after rearranging terms of Δ_l, we obtain

$$\Delta_l = (\sigma^4 + \kappa^2\delta^2 + 2\sigma^2\delta^2)y^2 - 2\delta(\kappa^2 + 2\sigma^2)y + \kappa^2 + 2\sigma^2$$

which is, again, quadratic in y. It's easy to calculate the discriminant of Δ_l, which is

$$-4(\kappa^2 + 2\sigma^2)\sigma^4 < 0.$$

Hence, $\Delta_l > 0$ for all y. In other words, we've shown that (15) always has two distinct positive roots since the constant term of the quadratic function $l(a)$ is

$$(\kappa - \delta)y - 1 = \frac{(\kappa - \delta)x}{\kappa\mu} - 1 > 0.$$

We take the smaller one as $a(\theta^*)$.

The last step is to prove that for such an $a(\theta^*)$ is indeed feasible in the sense that

$$e^{\theta^*} = e^{-\delta a(\theta^*)}\left(1 + \kappa a(\theta^*) - \frac{\sigma^2}{2}a(\theta^*)^2\right) > 0$$

(15) implies that

$$1 + \kappa a(\theta^*) - \frac{\sigma^2}{2}a(\theta^*)^2 = y\left(\kappa - \delta - (\sigma^2 + \kappa\delta)a(\theta^*) + \frac{\delta\sigma^2}{2}a(\theta^*)^2\right) = yg(a(\theta^*))$$

where $g(a)$ is as defined in the proof of Proposition 1. As discussed in the proof of Proposition 1, $g(a)$ has two distinct positive roots a_1 and a_2. Note that

$$l(a_1) = yg(a_1) + \frac{\sigma^2}{2}a_1^2 - \kappa a_1 - 1 = \frac{\sigma^2}{2}a_1^2 - \kappa a_1 - 1 < 0$$

as shown in the proof of Proposition 1. It follows that $a(\theta^*) < a_1$ since $a(\theta^*)$ is the smaller positive root of $l(a)$. Therefore, $g(a(\theta^*)) > 0$.

REFERENCES

Asmussen, S., and P. W. Glynn. 2007. *Stochastic simulation: Algorithms and analysis*. New York: Springer-Verlag.

Bucklew, J. A. 2004. *Introduction to rare event simulation*. New York: Springer-Verlag.

Dembo, A., and O. Zeitouni. 1998. *Large deviations techniques and applications*. 2nd ed. New York: Springer-Verlag.

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. 2009. Affine point processes and portfolio credit risk. Available via <http://www.stanford.edu/ dept/MSandE/people/faculty/giesecke/publications.html>.

Giesecke, K., and B. Kim. 2007. Estimating tranche spread 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. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Oksendal, B., and A. Sulem. 2007. *Applied stochastic control of jump diffusions*. 2nd ed. New York: Springer-Verlag.

Protter, P. E. 2005. *Stochastic integration and differential equations*. 2nd ed. New York: Springer-Verlag.

Zhang, X., P. W. Glynn, K. Giesecke, and J. Blanchet. 2009. Rare event simulation for a selft-exciting point process.

AUTHOR BIOGRAPHIES

XIAO-WEI ZHANG is currently a Ph.D. candidate in the Department of Management Science and Engineering at Stanford University. He graduated with a B.Sc. in Mathematics from Nankai University, China. His research interests include applied probability, rare event simulation and stochastic modeling in finance. His email address is <xiaoweiz@stanford.edu>.

PETER W. GLYNN received his Ph.D. in Operations Research from Stanford University in 1982. He then joined the faculty of the University of Wisconsin at Madison, where he held a joint appointment between the Industrial Engineering Department and Mathematics Research Center, and courtesy appointments in Computer Science and Mathematics. In 1987, he returned to Stanford, where he is now the Thomas Ford Professor of Engineering in the Department of Management Science and Engineering. He also has a courtesy appointment in the Department of Electrical Engineering and serves as Director of the Stanford Institute of Computational and Mathematical Engineering. He is a Fellow of INFORMS and the Institute of Mathematical Statistics and his research interests lie in computational probability, simulation, queueing theory, statistical inference for stochastic processes, and stochastic modeling.

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>.

JOSE BLANCHET is an Assistant Professor in the Department of Industrial Engineering and Operations Research at Columbia University. Jose holds a M.Sc. in Engineering- Economic Systems and Operations Research and a Ph.D. in Management Science and Engineering, both from Stanford University. He also holds two B.Sc. degrees: one in Actuarial Science and another one in Applied Mathematics from ITAM (Mexico). Jose worked for two years as an analyst in Protego Financial Advisors, a leading investment bank in Mexico. He has research interests in applied probability, computational finance, performance engineering, queueing theory, risk management, rare-event analysis, statistical inference, stochastic modelling, and simulation. His email address is <jose.blanchet@columbia.edu>.