A Wiener-Hopf Monte Carlo simulation technique for Lévy processes

A. E. Kyprianou, J. C. Pardo and K. van Schaik

Department of Mathematical Sciences, University of Bath

Motivation

Motivation

- **Lévy process.** A (one dimensional) process X with stationary and independent increments and cadlag paths (e.g. Brownian motion with drift, compound Poisson processes, stable processes amongst many others).

Motivation

- **Lévy process.** A (one dimensional) process X with stationary and independent increments and cadlag paths (e.g. Brownian motion with drift, compound Poisson processes, stable processes amongst many others).
- **Goal of this talk.** Present Monte Carlo simulation technique for problems that fundamentally depend on the joint distribution

$$P(X_t \in dx, \overline{X}_t \in dy)$$

where $\overline{X}_t := \sup_{s \leq t} X_s$.

Motivation

- **Lévy process.** A (one dimensional) process X with stationary and independent increments and cadlag paths (e.g. Brownian motion with drift, compound Poisson processes, stable processes amongst many others).
- **Goal of this talk.** Present Monte Carlo simulation technique for problems that fundamentally depend on the joint distribution

$$P(X_t \in dx, \overline{X}_t \in dy)$$

  where $\overline{X}_t := \sup_{s \le t} X_s$.
- Example: barrier options in Lévy market. Value of a European up-and-out barrier option with expiry date T and barrier b is of the form

$$e^{-rT}\mathbb{E}_s(f(S_T)\mathbf{1}_{\{\overline{S}_T \le b\}})$$

  where $S = \exp(X)$, $\overline{S}_T = \sup_{u \le T} S_u$ and f is some nice function.

Motivation

- **Lévy process.** A (one dimensional) process X with stationary and independent increments and cadlag paths (e.g. Brownian motion with drift, compound Poisson processes, stable processes amongst many others).
- **Goal of this talk.** Present Monte Carlo simulation technique for problems that fundamentally depend on the joint distribution

$$P(X_t \in dx, \overline{X}_t \in dy)$$

where $\overline{X}_t := \sup_{s \leq t} X_s$.

- Example: barrier options in Lévy market. Value of a European up-and-out barrier option with expiry date T and barrier b is of the form

$$e^{-rT}\mathbb{E}_s(f(S_T)\mathbf{1}_{\{\overline{S}_T \leq b\}})$$

where $S = \exp(X)$, $\overline{S}_T = \sup_{u \leq T} S_u$ and f is some nice function.

- Other motivations from queuing theory, population models etc.

Facts from Wiener-Hopf theory

Facts from Wiener-Hopf theory

- Recall characteristic exponent of X given by

$$\begin{aligned}\Psi(\theta) &:= -\frac{1}{t}\log E(e^{i\theta X_t}) \\ &= ai\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}}(1 - e^{i\theta x} + i\theta x\mathbf{1}_{\{|x|\le 1\}})\Pi(dx)\end{aligned}$$

where $a \in \mathbb{R}$, $\sigma \in \mathbb{R}$ and Π is a measure concentrated on $\mathbb{R}\backslash\{0\}$ satisfying $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$.

Facts from Wiener-Hopf theory

- Recall characteristic exponent of X given by

$$\begin{aligned}\Psi(\theta) &:= -\frac{1}{t}\log E(e^{i\theta X_t}) \\ &= ai\theta + \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}}(1 - e^{i\theta x} + i\theta x \mathbf{1}_{\{|x|\leq 1\}})\Pi(dx)\end{aligned}$$

  where $a \in \mathbb{R}$, $\sigma \in \mathbb{R}$ and Π is a measure concentrated on $\mathbb{R}\backslash\{0\}$ satisfying $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$.

- **Wiener-Hopf factorisation**: one can always decompose

$$q + \Psi(\theta) = \kappa^+(q, -i\theta) \times \kappa^-(q, i\theta)$$

  such that

$$E(e^{i\theta \overline{X}_{\mathbf{e}_q}}) = \frac{\kappa^+(q,0)}{\kappa^+(q,-i\theta)} \text{ and } E(e^{i\theta \underline{X}_{\mathbf{e}_q}}) = \frac{\kappa^-(q,0)}{\kappa^-(q,i\theta)}$$

  where $\mathbf{e}_q$ is an independent and exponentially distributed random variable with rate $q > 0$ and $\underline{X}_t := \inf_{s\leq t} X_s$. (Recall $\overline{X}_t := \sup_{s\leq t} X_s$.)

Facts from Wiener-Hopf theory

Facts from Wiener-Hopf theory

- In particular,

$$X_{\mathbf{e}_q} \stackrel{d}{=} S_q + I_q$$

where S_q is independent of I_q and they are respectively equal in distribution to $\overline{X}_{\mathbf{e}_q}$ and $\underline{X}_{\mathbf{e}_q}$.

Facts from Wiener-Hopf theory

- In particular,

$$X_{\mathbf{e}_q} \stackrel{d}{=} S_q + I_q$$

  where S_q is independent of I_q and they are respectively equal in distribution to $\overline{X}_{\mathbf{e}_q}$ and $\underline{X}_{\mathbf{e}_q}$.
- So:

$$(X_{\mathbf{e}_q}, \overline{X}_{\mathbf{e}_q}) \stackrel{d}{=} (S_q + I_q, S_q).$$

Facts from Wiener-Hopf theory

- In particular,
$$X_{\mathbf{e}_q} \stackrel{d}{=} S_q + I_q$$
where S_q is independent of I_q and they are respectively equal in distribution to $\overline{X}_{\mathbf{e}_q}$ and $\underline{X}_{\mathbf{e}_q}$.
- So:
$$(X_{\mathbf{e}_q}, \overline{X}_{\mathbf{e}_q}) \stackrel{d}{=} (S_q + I_q, S_q).$$
- **Q1**. For what LP's can we indeed sample from S_q and I_q?
Recent advances: there are many new examples of Lévy processes (with two-sided jumps) emerging for which sufficient analytical structure is in place in order to sample from the two distributions $\overline{X}_{\mathbf{e}_q}$ and $\underline{X}_{\mathbf{e}_q}$. ($\beta$-Lévy processes, Lamperti-stable processes, Hypergeometric Lévy processes, $\cdots$).

Facts from Wiener-Hopf theory

- In particular,

$$X_{\mathbf{e}_q} \stackrel{d}{=} S_q + I_q$$

  where S_q is independent of I_q and they are respectively equal in distribution to $\overline{X}_{\mathbf{e}_q}$ and $\underline{X}_{\mathbf{e}_q}$.

- So:

$$(X_{\mathbf{e}_q}, \overline{X}_{\mathbf{e}_q}) \stackrel{d}{=} (S_q + I_q, S_q).$$

- **Q1**. For what LP's can we indeed sample from S_q and I_q?

  Recent advances: there are many new examples of Lévy processes (with two-sided jumps) emerging for which sufficient analytical structure is in place in order to sample from the two distributions $\overline{X}_{\mathbf{e}_q}$ and $\underline{X}_{\mathbf{e}_q}$. ($\beta$-Lévy processes, Lamperti-stable processes, Hypergeometric Lévy processes, $\cdots$).

- **Q2**. How do we get from the random time $\mathbf{e}_q$ to the fixed time t we are after?

  Put i.i.d. exponentials 'after each other' to construct 'stochastic time grid' and make use of stat. indep. increments of X.

'Stochastic time grid'[1]

[1]Peter Carr has made use of this fact in a different way in the past in a finance setting and Ron Doney in a theoretical probabilistic setting.

'Stochastic time grid'[1]

- Suppose that $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \cdots$ is a sequence of i.i.d exp(1) distributed r.v.'s.

[1] Peter Carr has made use of this fact in a different way in the past in a finance setting and Ron Doney in a theoretical probabilistic setting.

'Stochastic time grid'[1]

- Suppose that $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \cdots$ is a sequence of i.i.d exp(1) distributed r.v.'s.
- Define the 'grid points' (with av. grid distance $1/\lambda$) for all $k \geq 0$:

$$\mathbf{g}(k, \lambda) := \sum_{i=1}^{k} \frac{1}{\lambda} \mathbf{e}^{(i)},$$

in particular for any $t > 0$ by the strong law of Large numbers

$$\mathbf{g}(n, n/t) = \sum_{i=1}^{n} \frac{t}{n} \mathbf{e}^{(i)} \stackrel{n \to \infty}{\longrightarrow} t \quad \text{a.s.}$$

[1] Peter Carr has made use of this fact in a different way in the past in a finance setting and Ron Doney in a theoretical probabilistic setting.

'Stochastic time grid'[1]

- Suppose that $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \cdots$ is a sequence of i.i.d exp(1) distributed r.v.'s.
- Define the 'grid points' (with av. grid distance $1/\lambda$) for all $k \geq 0$:

$$\mathbf{g}(k, \lambda) := \sum_{i=1}^{k} \frac{1}{\lambda} \mathbf{e}^{(i)},$$

in particular for any $t > 0$ by the strong law of Large numbers

$$\mathbf{g}(n, n/t) = \sum_{i=1}^{n} \frac{t}{n} \mathbf{e}^{(i)} \stackrel{n\to\infty}{\longrightarrow} t \quad \text{a.s.}$$

- Hence for a suitably large n, we have in distribution

$$(X_{\mathbf{g}(n,n/t)}, \overline{X}_{\mathbf{g}(n,n/t)}) \simeq (X_t, \overline{X}_t).$$

Indeed since t is not a jump time with probability 1, we have that $(X_{\mathbf{g}(n,n/t)}, \overline{X}_{\mathbf{g}(n,n/t)}) \to (X_t, \overline{X}_t)$ a.s. as $n \to \infty$.

[1] Peter Carr has made use of this fact in a different way in the past in a finance setting and Ron Doney in a theoretical probabilistic setting.

'Stochastic time grid'[1]

- Suppose that $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \cdots$ is a sequence of i.i.d exp(1) distributed r.v.'s.
- Define the 'grid points' (with av. grid distance $1/\lambda$) for all $k \geq 0$:

$$\mathbf{g}(k, \lambda) := \sum_{i=1}^{k} \frac{1}{\lambda} \mathbf{e}^{(i)},$$

  in particular for any $t > 0$ by the strong law of Large numbers

$$\mathbf{g}(n, n/t) = \sum_{i=1}^{n} \frac{t}{n} \mathbf{e}^{(i)} \stackrel{n \to \infty}{\longrightarrow} t \quad \text{a.s.}$$

- Hence for a suitably large n, we have in distribution

$$(X_{\mathbf{g}(n,n/t)}, \overline{X}_{\mathbf{g}(n,n/t)}) \simeq (X_t, \overline{X}_t).$$

  Indeed since t is not a jump time with probability 1, we have that $(X_{\mathbf{g}(n,n/t)}, \overline{X}_{\mathbf{g}(n,n/t)}) \to (X_t, \overline{X}_t)$ a.s. as $n \to \infty$.
- This + facts from W-H theory from previous slide yields main result:

[1] Peter Carr has made use of this fact in a different way in the past in a finance setting and Ron Doney in a theoretical probabilistic setting.

Main result

- **Theorem.** For all $n \in \{1, 2, \cdots\}$ and $\lambda > 0$,

$$(X_{\mathbf{g}(n,\lambda)}, \overline{X}_{\mathbf{g}(n,\lambda)}) \stackrel{d}{=} (V(n,\lambda), J(n,\lambda))$$

where

$$V(n,\lambda) := \sum_{j=1}^{n} \{S_\lambda^{(j)} + I_\lambda^{(j)}\} \text{ and } J(n,\lambda) := \bigvee_{i=0}^{n-1} \left(\sum_{j=1}^{i} \{S_\lambda^{(j)} + I_\lambda^{(j)}\} + S_\lambda^{(i+1)} \right).$$

Here:
- $\{S_\lambda^{(j)} : j \geq 1\}$ is an i.i.d. sequence of r.v.'s with common distribution equal to that of $\overline{X}_{\mathbf{e}_\lambda}$,
- $\{I_\lambda^{(j)} : j \geq 1\}$ is another i.i.d. sequence of r.v.'s with common distribution equal to that of $\underline{X}_{\mathbf{e}_\lambda}$.

Main result

- **Theorem.** For all $n \in \{1, 2, \cdots\}$ and $\lambda > 0$,

$$(X_{\mathbf{g}(n,\lambda)}, \overline{X}_{\mathbf{g}(n,\lambda)}) \stackrel{d}{=} (V(n,\lambda), J(n,\lambda))$$

  where

$$V(n,\lambda) := \sum_{j=1}^{n} \{S_\lambda^{(j)} + I_\lambda^{(j)}\} \text{ and } J(n,\lambda) := \bigvee_{i=0}^{n-1} \left(\sum_{j=1}^{i} \{S_\lambda^{(j)} + I_\lambda^{(j)}\} + S_\lambda^{(i+1)} \right).$$

  Here:
  - $\{S_\lambda^{(j)} : j \geq 1\}$ is an i.i.d. sequence of r.v.'s with common distribution equal to that of $\overline{X}_{\mathbf{e}_\lambda}$,
  - $\{I_\lambda^{(j)} : j \geq 1\}$ is another i.i.d. sequence of r.v.'s with common distribution equal to that of $\underline{X}_{\mathbf{e}_\lambda}$.

- With a.s. convergence from previous slide:
  Corollary. We have as $n \uparrow \infty$

$$(V(n, n/t), J(n, n/t)) \to (X_t, \overline{X}_t)$$

  where the convergence is understood in the distributional sense.

Example of implementation

- Setup Monte Carlo simulation:

$$\mathbb{E}(g(X_t, \overline{X}_t)) \simeq \frac{1}{m} \sum_{i=1}^{m} g(V^{(i)}(n, n/t), J^{(i)}(n, n/t)).$$

Example of implementation

- Setup Monte Carlo simulation:

$$\mathbb{E}(g(X_t, \overline{X}_t)) \simeq \frac{1}{m} \sum_{i=1}^{m} g(V^{(i)}(n, n/t), J^{(i)}(n, n/t)).$$

- Requirement: being able to sample from

$$I_{n/t} \stackrel{d}{=} \underline{X}_{\mathbf{e}_{n/t}} \quad \text{and} \quad S_{n/t} \stackrel{d}{=} \overline{X}_{\mathbf{e}_{n/t}}.$$

As mentioned before, recently new large families of LP's have occured for which this is possible.

Example of implementation

- Setup Monte Carlo simulation:

$$\mathbb{E}(g(X_t, \overline{X}_t)) \simeq \frac{1}{m} \sum_{i=1}^{m} g(V^{(i)}(n, n/t), J^{(i)}(n, n/t)).$$

- Requirement: being able to sample from

$$I_{n/t} \stackrel{d}{=} \underline{X}_{\mathbf{e}_{n/t}} \quad \text{and} \quad S_{n/t} \stackrel{d}{=} \overline{X}_{\mathbf{e}_{n/t}}.$$

  As mentioned before, recently new large families of LP's have occured for which this is possible.

- E.g. β-**family of LP's** by Kuznetsov (2009). Free to choose Gaussian part σ and drift part a; Lévy measure Π has density π given by

$$\pi(x) = c_1 \frac{e^{-\alpha_1 \beta_1 x}}{(1 - e^{-\beta_1 x})^{\lambda_1}} \mathbf{1}_{\{x>0\}} + c_2 \frac{e^{\alpha_2 \beta_2 x}}{(1 - e^{\beta_2 x})^{\lambda_2}} \mathbf{1}_{\{x<0\}}.$$

Example of implementation

- Setup Monte Carlo simulation:

$$\mathbb{E}(g(X_t, \overline{X}_t)) \simeq \frac{1}{m} \sum_{i=1}^{m} g(V^{(i)}(n, n/t), J^{(i)}(n, n/t)).$$

- Requirement: being able to sample from

$$I_{n/t} \stackrel{d}{=} \underline{X}_{\mathbf{e}_{n/t}} \quad \text{and} \quad S_{n/t} \stackrel{d}{=} \overline{X}_{\mathbf{e}_{n/t}}.$$

  As mentioned before, recently new large families of LP's have occured for which this is possible.

- E.g. β-**family of LP's** by Kuznetsov (2009). Free to choose Gaussian part σ and drift part a; Lévy measure Π has density π given by

$$\pi(x) = c_1 \frac{e^{-\alpha_1 \beta_1 x}}{(1 - e^{-\beta_1 x})^{\lambda_1}} \mathbf{1}_{\{x>0\}} + c_2 \frac{e^{\alpha_2 \beta_2 x}}{(1 - e^{\beta_2 x})^{\lambda_2}} \mathbf{1}_{\{x<0\}}.$$

- Note that the β-family of LP's has exponential moments (needed to work with risk neutral measures), there is asymmetry in the jump structure and locally jumps are stable-like (similarly to e.g. CGMY processes). Moreover we can have infinite or finite activity, bounded or unbounded path variation.

Example of implementation

- Kuznetsov uses that the characteristic exponent of X can be extended as a meromorphic function, together with analytical techniques, to identify the W-H factors and derive e.g.

$$P(\overline{X}_{\mathbf{e}_q} \in dx) = \left(\sum_{n \leq 0} k_n \zeta_n e^{\zeta_n x} \right) dx,$$

where the ζ_n's are (real) zeros of $z \mapsto q + \Psi(z)$ and have to be found numerically.

Example of implementation

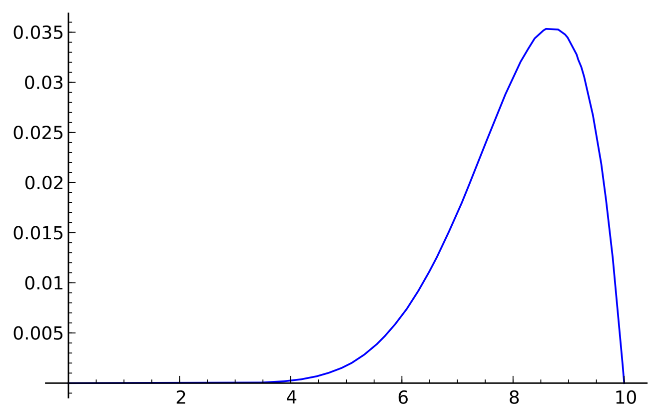
- Kuznetsov uses that the characteristic exponent of X can be extended as a meromorphic function, together with analytical techniques, to identify the W-H factors and derive e.g.

$$P(\overline{X}_{\mathbf{e}_q} \in dx) = \left(\sum_{n \leq 0} k_n \zeta_n e^{\zeta_n x} \right) dx,$$

  where the ζ_n's are (real) zeros of $z \mapsto q + \Psi(z)$ and have to be found numerically.

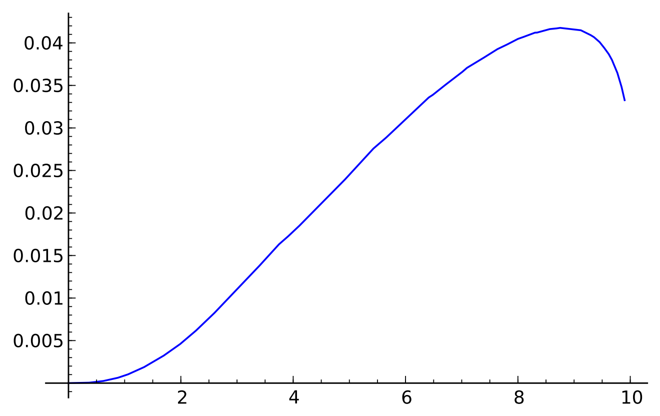
- A similar expression for $P(\underline{X}_{\mathbf{e}_q} \in dx)$

Simulated value function of European up-and-out call option with X from β-family



With Gaussian part (strike=5, barrier=10, T=1)

Simulated value function of European up-and-out call option with X from β-family



Irregular upwards (strike=5, barrier=10, T=1)

Advantages over standard random walk approach

- Standard random walk:
  - in general law of X_t not known, needs to be obtained by numerical Fourier inversion.
  - *always* produces an atom at 0 when simulating $\overline{X}_t$ (W-H MC method produces atom iff it is really present, i.e. iff X is irregular upwards).
  - well known bad performance when simulating $\overline{X}_t$ (misses excursions between grid points), W-H MC method performs significantly better in Brownian motion test case

W-H MC method vs. random walk: $P(\overline{X}_1 \leq z)$ where X is BM; $n =$ number of time steps (for r.w. $2n$ time steps)

		$z = 0.1$	$z = 0.2$	$z = 0.3$	$z = 0.4$	$z = 0.5$	$z = 1$	$z = 1.5$	$z = 2$
	exact	0.0797	0.1585	0.2358	0.3108	0.3829	0.6827	0.8664	0.9545
$n = 10$	w.h.	0.0828	0.1644	0.2447	0.3219	0.3955	0.6944	0.8700	0.9523
	error	3.88%	3.74%	3.75%	3.56%	3.28%	1.71%	0.41%	-0.23%
	r.w.	0.1886	0.2593	0.3315	0.4020	0.4689	0.7389	0.8951	0.9661
	error	136.76%	63.57%	40.56%	29.36%	22.44%	8.23%	3.32%	1.21%
$n = 100$	w.h.	0.0803	0.1592	0.2372	0.3125	0.3843	0.6852	0.8672	0.9546
	error	0.79%	0.41%	0.58%	0.52%	0.35%	0.36%	0.09%	0.01%
	r.w.	0.1122	0.1909	0.2675	0.3411	0.4116	0.7018	0.8764	0.9586
	error	40.90%	20.40%	13.45%	9.72%	7.48%	2.80%	1.16%	0.43%
$n = 1000$	w.h.	0.0792	0.1581	0.2357	0.3112	0.3837	0.6839	0.8665	0.9546
	error	-0.53%	-0.27%	-0.07%	0.12%	0.20%	0.17%	0.03%	0.00%
	r.w.	0.0899	0.1684	0.2456	0.3206	0.3925	0.6896	0.8699	0.9559
	error	12.91%	6.24%	4.16%	3.12%	2.50%	1.01%	0.41%	0.15%

Table 1: Computing $\mathbb{P}(\overline{X}_1 \leq z)$ for different values of z when X is a standard Brownian motion.

W-H MC method vs. random walk: $P(X_1 \leq z_1, \overline{X}_1 \geq z_2)$ where X is BM; 1000 time steps (for r.w. 2000 time steps)

		$z_2 = 0.1$	$z_2 = 0.3$	$z_2 = 0.5$	$z_2 = 1$
$z_1 = -2$	exact	0.0139	0.0047	0.0014	0.00003
	w.h.	0.0138	0.0046	0.0013	0.00003
	error	-0.93%	-1.93%	-1.33%	-5.27%
	r.w.	0.0128	0.0043	0.0012	0.00002
	error	-7.92%	-8.22%	-10.51%	-24.22%
$z_1 = -1$	exact	0.1151	0.0548	0.0228	0.0014
	w.h.	0.1147	0.0544	0.0225	0.0013
	error	-0.28%	-0.65%	-0.91%	-5.77%
	r.w.	0.1095	0.0515	0.0210	0.0012
	error	-4.87%	-6.12%	-7.54%	-14.36%
$z_1 = 0$	exact	0.4207	0.2743	0.1587	0.0228
	w.h.	0.4205	0.2738	0.1576	0.0223
	error	-0.06%	-0.18%	-0.68%	-2.02%
	r.w.	0.4101	0.2653	0.1518	0.0211
	error	-2.54%	-3.26%	-4.34%	-7.18%