Abstract

In this paper, we present an unbiased Monte Carlo estimator for lookback options in jump-diffusion models. Lookback options are difficult to price in jump-diffusion models, as their pay-off depends on the maximum of the share price over a particular time interval. In general, closed form solutions for prices of lookback options are not available but even simulating the pay-off of such an option is difficult. We show that after conditioning on jump times, the values of the share price just before and after the jumps and at maturity, we recover a finite number of Brownian bridges. The law of the maximum of a Brownian Bridge is well-known and we show in this paper that it can be simulated from. This realization allows us to derive an unbiased Monte Carlo estimator for lookback options in jump-diffusion models.
The contribution of this paper is an unbiased estimator for the prices of lookback options in jump-diffusion models. After the seminal papers introducing the Black-Scholes model, [3] and [10], a lot of practical and theoretical interest has been devoted to addressing the well-known shortcomings of the assumptions underlying the model, one of which postulates that stock prices follow a geometric Brownian motion. Empirically, it has been observed that stock prices do exhibit jumps, which a continuous process such as Brownian motion cannot account for. One way of addressing this phenomenon is to model the stock price process using a jump-diffusion model, e.g. [9], [7] and [8]. The properties of jump-diffusion processes relevant for this paper are recalled in section 2. As jump-diffusion processes are less tractable than diffusions, there is a need for unbiased numerical methods to price options, exotic options in particular.

One example of a class of exotic options which are difficult to price analytically in jump-diffusion models are lookback options, e.g. [8]. We introduce lookback options in section 2 and summarize the difficulties in pricing lookback options at the beginning of section 3.

In section 3, we introduce an estimator which allows to derive unbiased estimates of prices of lookback options. The estimator is discussed in section 3 and we derive a formula for the variance of the estimator. Finally, we present an unbiased estimator of the variance of the estimator of prices of lookback options. An algorithm to compute the estimator of prices of lookback options and the estimator of the variance of the estimator of prices of lookback options is presented in section 4.

The Kou model, [7], [8], a particular jump-diffusion model, is introduced in section 5. The algorithm introduced in section 4 is applied to the problem of pricing a lookback option in the Kou model and the results are compared to those produced by a previous, biased approach to pricing lookback options in a jump-diffusion model, which is summarized at the beginning of section 3.

Section 6 concludes, describes shortcomings of the methodology introduced and future areas of research.

2 Jump-Diffusion Processes and Lookback options

In this section, jump-diffusion processes and lookback options are introduced.

2.1 Jump-Diffusion Processes

We fix a time horizon $T^+$ and assume that we deal with a filtered probability space $(\Omega, \mathcal{F}_{T^+}, \mathbb{F}, \mathbb{P}^*)$. On this filtered probability space, we now introduce a jump-diffusion process.

Definition 2.1 Jump-diffusion process:
Let $(N_t)$ be a Poisson Process with intensity $\lambda$ and waiting times $\tau$, $(W_t)$ a Brownian Motion and $\{Y_i\}$ independent, identically distributed random variables with distribution $f$, then

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$$

is a jump-diffusion process, where $\sigma \in \mathbb{R}^+$ and $\gamma \in \mathbb{R}$.

Remark 2.1 Let $b = \int_{\mathbb{R}} x 1_{|x| \leq 1} \nu(dx)$, then it is easy to see that $X_t$ is a Lévy process with characteristic triplet $(\gamma + b, \sigma, \nu)$, where $\nu(A) = \lambda f(A)$, $A \in \mathcal{B}(\mathbb{R})$.

It is assumed that $N_t$, $W_t$ and the $\{Y_i\}$ are independent and we assume that the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T^+]}$ is the $\mathbb{P}^*$-augmentation of the natural filtration of the noise process, so $\mathcal{F}_t = \mathcal{F}_t^\mathbb{P}$.
holds for \( t \in [0, T^+] \), where \( \mathcal{F}_t^n = \sigma(N_u, W_u, u \leq t, Y_i, i \leq N_t) \).

We are now in a position to introduce the stock price process on \((\Omega, \mathcal{F}_T^*, \mathbb{F}, \mathbb{P}^*)\).

**Definition 2.2** Stock price process:
Let
\[
\gamma = r - \frac{1}{2} \sigma^2 - \lambda \int_{-\infty}^{\infty} (e^x - 1) f(dx)
\]
and set
\[
S_t = S_0 \exp \left( X_t \right)
\]
(1)
Then \( S \) is the stock price process.

From the definition of \( \gamma \), the following lemma follows:

**Lemma 2.1** \( \frac{S_t}{\exp(rt)} \) is a martingale.

**Proof.** It was stated in Remark 2.1 that \( X_t \) is a Lévy process. The definition of \( \gamma \) yields that
\[
\mathbb{E}_{\mathbb{P}^*}[\exp(X_t)] = \exp(rt)
\]
It now follows from Proposition 3.17 in [5] that \( \left( \frac{\exp(X_t)}{\exp(rt)} \right)_{t \geq 0} \) is a martingale, hence \( \left( \frac{S_t}{\exp(rt)} \right)_{t \geq 0} \) is a martingale.

In what follows, we assume that \( \mathbb{P}^* \) is the measure under which prices of options are computed. This assumption results in the following definition.

**Definition 2.3** Prices of European Options:
The price of a possibly path-dependent European option, with pay-off \( \psi_S(T) \) at maturity \( T \), is given by:
\[
\psi_S(t) = \exp \left( -r(T - t) \right) \mathbb{E}_{\mathbb{P}^*}[\psi_S(T) | \mathcal{F}_t]
\]
(2)

**Remark 2.2** In financial mathematics, the share price is usually modeled on a probability space \((\Omega, \mathcal{F}_T^*, \mathbb{F}, \mathbb{P})\), where \( \mathbb{P} \) is the real-world probability measure. Consequently, we seek to find a probability measure, usually denoted by \( \mathbb{P}^* \), so that \( \mathbb{P}^* \) is equivalent to \( \mathbb{P} \) and discounted share prices follow a martingale under \( \mathbb{P}^* \). For jump-diffusion models, it is well-known that due to the presence of jumps, such a measure \( \mathbb{P}^* \) is not unique. Kou, [7], for example uses a rational expectations argument to arrive at a martingale measure \( \mathbb{P}^* \). However, this paper is concerned with introducing a particular numerical method, and hence we assume that we are given a martingale measure \( \mathbb{P}^* \), under which we price options.

### 2.2 Lookback Options

We firstly give a definition of lookback options.

**Definition 2.4** Lookback Options:
A lookback call option, maturing at time \( T \), is characterized by the following pay-off at time \( T \)
\[
LC(T) = S_T - \min \left( A, \min_{0 \leq t \leq T} S_t \right)
\]
(3)
where \( A \in \mathbb{R}^+ \). Similarly, a Lookback Put option, maturing at time \( T \), is characterized by the following pay-off at \( T \)
\[
LP(T) = \max \left( B, \max_{0 \leq t \leq T} S_t \right) - S_T
\]
(4)
where \( B \in \mathbb{R}^+ \).

The following corollary follows immediately from Definition 2.3.
Corollary 2.1 Prices of Lookback Options:
The prices of Lookback Call and Put options, $LC(0)$ and $LP(0)$ respectively, are given by

\begin{align}
LC(0) &= \mathbb{E}_P^* \left[ \exp(-rT)LC(T) \right] \\
LP(0) &= \mathbb{E}_P^* \left[ \exp(-rT)LP(T) \right]
\end{align}

Remark 2.3 It is worth noting that in the Black-Scholes model, [3], [10], closed form solutions for prices of lookback call and put options are available, e.g. [12]. However, due to the additional complexity of jump-diffusion models, closed-form solutions for prices of lookback options need not be available, which suggests the need for numerical methods allowing to estimate option prices without bias.

3 An unbiased estimator for Lookback Options

In this section, we firstly present a previous, biased approach to pricing lookback options, then introduce our estimator for lookback options, show that it is unbiased and lastly investigate its variance. We only produce results for lookback put options, the corresponding results for lookback call options are left to the reader.

We recall that the price of a lookback option is given by the following formula:

\begin{align}
LP(0) &= \mathbb{E}_P^* \left[ \exp(-rT) \max_{0 < t < T} S_t - S_0 \right] \\
&= \mathbb{E}_P^* \left[ \exp(-rT) \max_{0 < t < T} S_t \right] - S_0
\end{align}

It is clear that the estimation problem is reduced to estimating

\begin{align}
\mathbb{E}_P^* \left[ \exp(-rT) \max_{0 < t < T} S_t \right]
\end{align}

The difficulties encountered when trying to price lookback options in jump-diffusion models are the following: The law of $S_t$ may be difficult to derive, depending on the underlying jump-diffusion process. This makes it difficult to price options depending on

\begin{align}
\max_{0 < t < T} S_t
\end{align}

and in general, one cannot expect to be able to arrive at closed-form expressions for prices of lookback options. For a special case where closed-form solutions in terms of special functions can be computed, see [8]. Even simulating random variates

\begin{align}
\max_{0 < t < T} S_t
\end{align}

is not possible in general, due to the intractability of jump-diffusion processes, making the implementation of Monte Carlo methods difficult.

A biased Monte Carlo method, which approximates the random variates

\begin{align}
\max_{0 < t < T} S_t
\end{align}

by variates easier to simulate, is introduced in Section 3.1. An unbiased estimator is presented and analyzed in the subsequent sections.
3.1 A previous approach to the problem

In this section, we present a previous approach to pricing lookback options in jump-diffusion models, which is for example used in [8]. We also use this opportunity to present some basic Monte Carlo results in the context of our option pricing problem. We now introduce a biased estimator for the price of a lookback put option. The function

$$\exp(-rT) \max_{0 < t < T} S_t$$

is approximated using the following function, which computes the maximum over a finite number of points, \(m_2\), instead of an infinite number of points as in the original formulation:

$$\hat{f}(S_{t_1}, \ldots, S_{t_{m_2}}) = \exp(-rT) \max \left( B, \max_{i = 1, \ldots, m_2} S_{t_i} \right)$$

and set

$$O_{m_2} = \hat{f}(S_{t_1}, \ldots, S_{t_{m_2}})$$

\((t_i)_{i=1}^{m_2}\) is a partition of \([0, T]\) satisfying \(t_0 < t_1 < \ldots < t_{m_2-1} < t_{m_2} = T\). Furthermore, let \(O_{m_2i}, i = 1, \ldots, m_1\) be independent copies of \(O_{m_2}\), so \(O_{m_2i}\) are independent, identically distributed. We are now in a position to introduce the biased estimator.

**Definition 3.1** Let

$$\bar{O}_{m_2} = \frac{1}{m_1} \sum_{i=1}^{m_1} O_{m_2i}$$

Firstly, it can be sown that \(\bar{O}_{m_2}\) is an unbiased estimator of \(E[O_{m_2}]\) and we can compute the variance of \(\bar{O}_{m_2}\).

**Lemma 3.1** Let \(O_{m_2}\) and \(\bar{O}_{m_2}\) be as defined above. Then

1. \(\bar{O}_{m_2}\) is an unbiased estimator of \(E[O_{m_2}]\)
2. the variance of \(\bar{O}_{m_2}\) is \(\frac{\text{Var}(O_{m_2})}{m_1}\)

**Proof.**

1. follows as

$$E[\bar{O}_{m_2}] = \frac{1}{m_1} \sum_{i=1}^{m_1} E[O_{m_2i}] = \frac{1}{m_1} \sum_{i=1}^{m_1} E[O_{m_2}] = E[O_{m_2}]$$

The above follows as \(O_{m_2i}, i = 1, \ldots, m_1\) are identically distributed.

2. follows as

$$\text{Var}(\bar{O}_{m_2}) = \frac{1}{m_1^2} \sum_{i=1}^{m_1} \text{Var}(O_{m_2i}) = \frac{1}{m_1^2} \sum_{i=1}^{m_1} \text{Var}(O_{m_2}) = \frac{\text{Var}(O_{m_2})}{m_1}$$

The above follows as \(O_{m_2i}, i = 1, \ldots, m_1\) are independent, identically distributed. 

The next lemma shows that \(\bar{O}_{m_2}\) is a biased estimator of

$$E \left[ \exp(-rT) \max_{0 < t < T} S_t \right]$$

**Lemma 3.2** \(\bar{O}_{m_2}\) is a biased estimator of

$$E \left[ \exp(-rT) \max_{0 < t < T} S_t \right]$$
Proof.

\[ \mathbb{E} \left[ \bar{O}_{m_2} \right] = \mathbb{E} [ O_{m_2} ] \neq \mathbb{E} \left[ \exp(-rT) \max_{0 \leq t \leq T} S_t \right] \]

The next lemma produces an unbiased estimator of \( \frac{\text{Var}(O_{m_2})}{m_1} \), which is important, as it is used for example when deriving confidence intervals for \( \bar{O}_{m_2} \).

**Lemma 3.3**

\[ V_{O_{m_2}} = \frac{1}{m_1 (m_1 - 1)} \sum_{i=1}^{m_1} (O_{m_2i} - \bar{O}_{m_2})^2 \]

Then \( V_{O_{m_2}} \) is an unbiased estimate of \( \frac{\text{Var}(O_{m_2})}{m_1} \).

**Proof.**

\[ (O_{m_2i} - \bar{O}_{m_2})^2 = O_{m_2i}^2 - 2O_{m_2i} \sum_{i=1}^{m_1} \frac{O_{m_2i}}{m_1} + \left( \sum_{i=1}^{m_1} \frac{O_{m_2i}}{m_1} \right)^2 \]

Hence

\[ \sum_{i=1}^{m_1} (O_{m_2i} - \bar{O}_{m_2})^2 = \sum_{i=1}^{m_1} O_{m_2i}^2 - 2 \sum_{i=1}^{m_1} O_{m_2i} \sum_{i=1}^{m_1} \frac{O_{m_2i}}{m_1} + \sum_{i=1}^{m_1} \sum_{i=1}^{m_1} \frac{O_{m_2i}}{m_1} \]

= \sum_{i=1}^{m_1} O_{m_2i}^2 - \sum_{i=1}^{m_1} O_{m_2i} \sum_{i=1}^{m_1} \frac{O_{m_2i}}{m_1}

Consequently

\[ \mathbb{E} \left[ \sum_{i=1}^{m_1} (O_{m_2i} - \bar{O}_{m_2})^2 \right] = \sum_{i=1}^{m_1} \mathbb{E} \left[ O_{m_2i}^2 \right] - \frac{1}{m_1} \mathbb{E} \left[ \sum_{i=1}^{m_1} O_{m_2i}^2 + \sum_{i=1}^{m_1} \sum_{k \neq i} O_{m_2i} O_{m_2k} \right] \]

= \( m_1 \mathbb{E} \left[ O_{m_2i}^2 \right] - \frac{1}{m_1} \left( \sum_{i=1}^{m_1} \mathbb{E} \left[ O_{m_2i}^2 \right] + \sum_{i=1}^{m_1} \sum_{k \neq i} \mathbb{E} \left[ O_{m_2i} \right] \mathbb{E} \left[ O_{m_2k} \right] \right) \)

= \( (m_1 - 1) \left( \mathbb{E} \left[ O_{m_2i}^2 \right] + \left( \mathbb{E} \left[ O_{m_2i} \right] \right)^2 \right) \)

It now follows that

\[ \mathbb{E} \left[ V_{O_{m_2}} \right] = \frac{\mathbb{E} \left[ O_{m_2}^2 \right] + \left( \mathbb{E} \left[ O_{m_2} \right] \right)^2}{m_1} = \frac{\text{Var} \left( O_{m_2} \right)}{m_1} \]

as required. \( \square \)

Finally we present an algorithm to compute \( \bar{O}_{m_2} \) and \( V_{O_{m_2}} \). This algorithm is used in section 5.2 to estimate prices of lookback put options in the Kou model, \([7],[8] \), and to estimate variances of the estimators. By convention, \( N_{t_0} = 0, P(a) \) denotes a Poisson variate with parameter \( a \) and \( N(\mu, \sigma^2) \) a Normal variate with parameters \( \mu, \sigma^2 \).

**Algorithm 3.1** Algorithm to estimate \( \bar{O}_{m_2} \) and \( V_{O_{m_2}} \)

1. for \( i = 1 : m_1 \)
   
   (a) \( k = 1 \)
   
   (b) for \( j = 1 : m_2 \)
      
      i. simulate \( N_{t_j} \) by setting
      
      \[ N_{t_j} = N_{t_{j-1}} + P(\lambda (t_j - t_{j-1})) \]
ii. for \( l = 1 : (N_{t_{j,i}} - N_{t_{j-1,i}}) 
\)
   A. simulate \( Y_{ki} \) from distribution \( f \)
   B. \( k = k + 1 \)
   iii. end
   iv. \( W_{t_{j,i}} = W_{t_{j-1,i}} + N(0, t_{j} - t_{j-1}) \)
   v. \( S_{t_{j,i}} = S_{0} \exp \left( \gamma t_{j} + \sigma W_{t_{j,i}} + \sum_{k=1}^{N_{t_{j,i}}} Y_{ki} \right) \)
(c) end
(d) \( O_{m2i} = \exp(-rT) \max_{0 < i \leq T} \left( B, \max_{j=1,\ldots,m2} \left( S_{t_{j,i}} \right) \right) \)

2. end
3. \( \bar{O}_{m2} = \sum_{i=1}^{m1} \frac{O_{m2i}}{m1} \)
4. \( V_{O_{m2}} = \frac{1}{m1 (m1 - 1)} \sum_{i=1}^{m1} \left( O_{m2i} - \bar{O}_{m2} \right)^2 \)

In section 5.2, we will estimate prices of lookback put options and the variances of the estimators, for different values of \( m_1 \) and \( m_2 \).

3.2 The Monte Carlo Estimator

We recall that the estimation problem has been reduced to estimating 
\[ \mathbb{E}_{\mathbb{P}} \left[ \exp(-rT) \max_{0 < t < T} S_{t} \right] \] .

It is firstly noted that we can rewrite 
\[ \left[ \exp(-rT) \max_{0 < t < T} S_{t} \right] \]
as a function, say \( g \), of the random variables \( N_T, \tau_1, \ldots, \tau_{N_T}, Y_1, \ldots, Y_{N_T}, W_{\tau_1}, \ldots, W_{\tau_{N_T}}, M_1, \ldots, M_{N_T}, M_{N_T+1} \) as follows:
\[ g \left( N_T, \tau_1, \ldots, \tau_{N_T}, Y_1, \ldots, Y_{N_T}, W_{\tau_1}, \ldots, W_{\tau_{N_T}}, M_1, \ldots, M_{N_T}, M_{N_T+1} \right) = \exp(-rT) \max_{i=1,\ldots,N_T+1} M_i \]
where
\[ M_i = \max_{\tau_{i-1} < \tau_i \leq \tau} S_{\tau} \]
and \( \tau_0 = 0 \) and \( \tau_{N_T+1} = T \). To introduce the estimator, we need the following notation: Let \( N_{T,i}, \tau_{i,1}, \ldots, \tau_{N_{T},i,i}, Y_{i,1}, \ldots, Y_{i,N_{T},i,i}, W_{\tau_{i,1,i}}, \ldots, W_{\tau_{N_{T},i,i}}, i = 1, \ldots, m_1 \) be independent copies of \( N_T, \tau_1, \ldots, \tau_{N_T}, Y_1, \ldots, Y_{N_T}, W_{\tau_1}, \ldots, W_{\tau_{N_T}} \). We furthermore define the vectors \( \tau_i = (\tau_{1,i}, \ldots, \tau_{N_{T},i,i}) \),
Let $Y_i = (Y_{1,i}, \ldots, Y_{N_T,i})$ and $W_i = (W_{\tau_1,i}, \ldots, W_{\tau_{N_T,i}})$; clearly, the $N_T, \tau_i, Y_i, W_i$ are independently, identically distributed (iid). Furthermore, we introduce $M_{1,i,j}, \ldots, M_{N_T,i,j}, M_{N_T+1,i,j}$, $j = 1, \ldots, m_2$ which are independent copies of $M_{1,i}, M_{N_T,i}, M_{N_T+1,i}, i = 1, \ldots, m_1$. We define the matrix $M_{i,j} = M_{1,i,j}, \ldots, M_{N_T,i,j}, M_{N_T+1,i,j}, i = 1, \ldots, m_1, j = 1, \ldots, m_2$. Clearly for fixed $i$, $M_{i,j}, j = 1, \ldots, m_2$ are iid and for $i \neq k$, $M_{i,j}$ and $M_{k,j}$, $j = 1, \ldots, m_2, l = 1, \ldots, m_2$ are iid. We consequently motivate the estimator as follows: We cannot immediately simulate the random variable

$$ \max_{0 \leq t \leq T} S_t $$

for $S_t$ given by (1), which would however be possible for $\lambda = 0$; this observation motivates our estimator.

The following lemma shows a useful way of rewriting

$$ \mathbb{E}_\mathbb{P}_\tau \left[ \exp(-rT) \max_{0 \leq t \leq T} S_t \right] $$

All processes are defined as in section 2.1.

**Lemma 3.4** Let $(S_t)_{t \in [0,T]}$ be the stock price process and $T \leq T^*$. Then the price of a lookback put option can be represented as follows:

$$ LP_0 = \mathbb{E}_\mathbb{P}_\tau \left[ \exp(-rT) \max_{0 \leq t \leq T} B_s \right] $$

where $B_s = \exp(\sigma X_{\tau_1})$, $X_{\tau_1} = \sigma W_{\tau_1} + \gamma \tau_1$, $\tau_1 = \inf\{t \geq 0 : X_t = 0\}$, and $\tau_1 \in \tau_i, i = 1, \ldots, N_T$ denote the jump times of $(N_t)$, $\tau_0 = 0$ and $\tau_{N_T+1} = T$.

**Proof.** The proof follows immediately from the Tower Property.

The following lemmas show how to simulate $M_i$ conditional on $\mathcal{F}^\tau$. The first lemma is based on reasoning in [5], p. 177, but included for completeness.

**Lemma 3.5** Let $M_i$ and $\mathcal{F}^\tau$ be defined as in Lemma 3.4. Then

1. The $M_i$ are mutually independent

2. $\mathbb{P}^\tau (M_i \leq b | \mathcal{F}^\tau) = 1 - \exp\left(-\frac{2(-x-a)(x-a)}{\sigma^2}\right)$ where $l = \tau_i - \tau_{i-1}, z = X(\tau_i), x = X(\tau_{i-1})$ and $a = \ln \left( \frac{b}{S_0} \right)$

**Proof.** We firstly make the observation that because we have conditioned on the jumps, $M_i$ only depends on the Brownian motion but does not jump. We now proceed with the proof.

1. follows immediately from the Markov Property of Brownian Motion.

The argument for 2. is based on the reasoning in [5], p. 177, and shows how to recover the well-known law of the maximum of a Brownian Bridge with drift, e.g. [2]:

$$ \mathbb{P}^\tau (M_i \leq b | \mathcal{F}^\tau) = \mathbb{P}^\tau (M_i \leq b | X_{\tau_{i-1}}, X_{\tau_i}) $$

$$ = \mathbb{P}^\tau \left( \max_{\tau_{i-1} \leq t < \tau_i} S_t \leq b | X_{\tau_{i-1}}, X_{\tau_i} \right) = \mathbb{P}^\tau \left( \max_{\tau_{i-1} \leq t < \tau_i} \exp(X_t) \leq \frac{b}{S_0} | X_{\tau_{i-1}}, X_{\tau_i} \right) $$

$$ = \mathbb{P}^\tau \left( \max_{\tau_{i-1} \leq t < \tau_i} X_t \leq \ln \left( \frac{b}{S_0} \right) | X_{\tau_{i-1}}, X_{\tau_i} \right) $$

$$ = \mathbb{P}^\tau X_{\tau_{i-1}} \left( \sup_{0 \leq t < \tau_{i-1}} \sigma \hat{W}_t + \gamma t \leq a | \sigma \hat{W}_{\tau_{i-1}}, \gamma (\tau_i - \tau_{i-1}) = X_{\tau_i} \right) $$

$$ = \left\{ \begin{array}{ll}
1 - \exp \left( -\frac{2(-x-a)(x-a)}{\sigma^2} \right) & \text{for } a \geq \max(x,z) \\
0 & \text{otherwise}
\end{array} \right. $$


8

**LOOKBACK OPTIONS IN JUMP-DIFFUSION MODELS**
where \( \hat{W} \) is a new Brownian motion and \( \mathbb{P}^* \) a probability measure under which \( \sigma \hat{W}_0 = x \). Equation (8) follows from the Markov Property of Brownian Motion, (9) follows from the fact that no jumps occur on \([\tau_{i-1}, \tau_i)\) and (10) is taken from [2].

The next lemma shows that one can simulate from the distribution defined under Lemma 3.5, 2.

**Lemma 3.6** Let \( M_t \) and \( \mathcal{F}^* \) be as in Lemma 3.4. Then we can simulate from the distribution of \( M_t \), conditional on \( \mathcal{F}^* \).

**Proof.** Let \( u \) be uniformly distributed in \([0, 1]\), then

\[
u = \mathbb{P}^* (M_t \leq b | \mathcal{F}^*)
\]

\[
\iff u = \mathbb{P}_x^* \left( \sup_{0 \leq t \leq l} \sigma \hat{W}_t + \gamma t \leq \ln \left( \frac{b}{\delta_0} \right) | \sigma \hat{W}_t + \gamma t = z \right)
\]

\[
\iff u = 1 - \exp \left( -\frac{2(z - a)(x - a)}{l\sigma^2} \right)
\]

where \( x = X_{\tau_{i-1}} \), \( a = \ln \left( \frac{b}{\delta_0} \right) \), \( z = X_{\tau_i} \), \( l = \tau_i - \tau_{i-1} \)

\[
\begin{align*}
  u &= 1 - \exp \left( -\frac{2(z - a)(x - a)}{l\sigma^2} \right) \\
  \iff \ln (1 - u) &= -2(z - a)(x - a) \frac{l\sigma^2}{2} \\
  \iff -l\sigma^2 \ln(1 - u) &= 2zx - az - ax + a^2 \\
  \iff a^2 + a(-z - x) + zx + \frac{l\sigma^2 \ln(1 - u)}{2} &= 0
\end{align*}
\]

hence

\[
a_{1,2} = \frac{z + x \pm \sqrt{(z + x)^2 - 4zx + l\sigma^2 \ln(1 - u)}}{2}
\]

\[
\iff a_{1,2} = \frac{(z + x) + \sqrt{(z - x)^2 - 2l\sigma^2 \ln(1 - u)}}{2}
\]

but \( a \geq \max(z, x) \), hence

\[
a = \frac{(z + x) + \sqrt{(z - x)^2 - 2l\sigma^2 \ln(1 - u)}}{2}
\]

so

\[
b = S_0 \exp(a)
\]

The intuition behind the estimator is the following. Starting from expression (7), we initially simulate \( N_{T,1}, \tau_{1,1}, \ldots, \tau_{N_{T,1},1}, Y_{1,1}, \ldots, Y_{N_{T,1},1}, W_{\tau_{1,1},1}, \ldots, W_{\tau_{N_{T,1},1},1} \), then estimate the inner conditional expectation for the simulated values \( N_{T,1}, \tau_{1,1}, \ldots, \tau_{N_{T,1},1}, Y_{1,1}, \ldots, Y_{N_{T,1},1}, W_{\tau_{1,1},1}, \ldots, W_{\tau_{N_{T,1},1},1} \). Consequently, the set of variates \( N_{T,2}, \tau_{2,2}, \ldots, \tau_{N_{T,2},2}, Y_{1,2}, \ldots, Y_{N_{T,2},2}, W_{\tau_{1,2},2}, \ldots, W_{\tau_{N_{T,2},2},2} \) is generated, the corresponding conditional expectation is estimated and this procedure is repeated. Finally, the average over the estimates of the conditional expectations corresponding to the different sets of \( N_{T,1}, \tau_{1,1}, \ldots, \tau_{N_{T}}, Y_{1,1}, \ldots, Y_{N_{T}}, W_{\tau_{1,1}}, \ldots, W_{\tau_{N_{T}}} \) is computed. This is formalized as follows:

**Definition 3.2** The estimator for

\[
\mathbb{E}_{\mathbb{P}} \left[ \exp(-rT) \max_{0 \leq t \leq T} S_t \right]
\]
Lookback Options in Jump-diffusion models

is

\[ I = \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) \]  
(13)

The first sum, running over index \(i\), corresponds to the outer expectation, the inner sum, running over index \(j\), corresponds to the inner expectation. For each \(i\), we simulate a set of variates \(N_{T,i}, \tau_i, Y_i, W_i\). For these particular variates, we estimate \(E_{P^*} [g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) | N_{T,i}, \tau_i, Y_i, W_i] \) using

\[ \frac{1}{m_2} \sum_{j=1}^{m_2} g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) \]

Finally, the outer loop \(N_{T,i}, \tau_i, Y_i, W_i\) averages over the estimates of the conditional expectations.

To summarize, the problem associated with simulating \(\max_{0 < t < T} S_t\) is solved in the following way: Due to the presence of jumps, it is not possible to simulate \(\max_{0 < t < T} S_t\) immediately, however this would be possible for \(\lambda = 0\). By initially simulating \(N_{T,i}, \tau_i, Y_i, W_i\) and consequently conditioning on these values, we are left to simulate \(M_{ij}\), which are known not to jump. Hence by simulating \(N_{T,i}, \tau_i, Y_i, W_i\) and conditioning on these variates, we are left to simulate variates only depending on the supremum of a Brownian Bridge with drift, which we can do by Lemma 3.6.

### 3.3 The estimator is unbiased

The contribution of this paper is to introduce an unbiased estimator for lookback options in jump-diffusion models.

Previous attempts to price lookback options in jump-diffusion models, as for example described in section 3.1, introduce a bias in the estimates of option prices. The following proposition shows that the estimator defined by (13) is unbiased.

**Proposition 3.1** The estimator (13) is unbiased.

**Proof.**

\[
E_{P^*} [I] = E_{P^*} \left[ \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) \right]
\]

\[
= \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} E_{P^*} [g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j})]
\]

\[
= \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} E_{P^*} \left[ \exp(-rT) \max_{0 < t < T} S_t \right]
\]

\[
= E_{P^*} \left[ \exp(-rT) \max_{0 < t < T} S_t \right]
\]

which proves the required. \(\square\)

### 3.4 Convergence of the variance of the estimator

In this section we investigate the order of magnitude of the variance of estimator (13). The variance of estimator (13) needs to be investigated, as it entails a double summation and the variables \(g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j})\) are not all independent.

The proofs of Lemma 3.8 and Proposition 3.2 make use of the following well-known result.
Lemma 3.7 Let \( Y, X_1, X_2, \ldots, X_n, n \geq 1 \) be random variables. Then

\[
E[Y | X_1, X_2, \ldots, X_n]
\]

is a function of \( X_1, X_2, \ldots, X_n, n \geq 1 \).

Remark 3.1 Lemma 3.7 only gives the existence of such a function but no explicit formula.

Lemma 3.8 The variance of estimator (13) is given by the following formula

\[
\text{Var}(I) = \frac{\text{Var}\left(f_1(N_T, \tau, Y, W)\right)}{m_1} + \frac{1}{m_1 m_2} \left( E[f_2(N_T, \tau, Y, W)] - E\left[(f_1(N_T, \tau, Y, W))^2\right] \right)
\]

Proof. See appendix A.

The following result follows easily from Lemma 3.8.

Lemma 3.9 1. For fixed \( m_2 \), the variance of estimator (13) is order of magnitude \( O\left(m_1^{-1}\right)\)

2. For fixed \( m_1 \), the variance is not order of magnitude \( O\left(m_2^{-1}\right)\)

Proof. Immediate from Lemma 3.8.

The algorithm 4.1 shows how to compute \( J \), the unbiased estimator of the variance of \( I \). Proposition 3.2 is used in section 5.2, when variances of the estimator \( I \) are estimated.

4 Algorithm to estimate prices of lookback options without bias

In this section, we present an algorithm which allows to compute estimators (13) and (14). The following result is used to simulate the arrival of the jumps and stems from [13], Theorem 5.2.
Lemma 4.1 Given that $N_t = n$, the arrival times $\tau_1, \ldots, \tau_n$ have the same distribution as the order statistics corresponding to $n$ independent random variables uniformly distributed on the interval $(0, t)$.

Proof. See [13].

The formulas (11) and (12), produced in Lemma 3.6, are also used in the algorithm. In the following, $N(\mu, \sigma^2)$ denotes a random variable with mean $\mu$ and variance $\sigma^2$.

Algorithm 4.1 Algorithm to compute estimators (13) and (14)

1. for $i = 1 : m_1$ (Outer loop)
   
   a. simulate $N_{T,i}$ as a Poisson random variate with parameter $\lambda T$
   b. simulate $\tau_{1,i}, \ldots, \tau_{N_{T,i},i}$ by generating $N_{T,i}$ independent random variates, uniformly distributed in $[0, T]$, $U_{1,i}, \ldots, U_{N_{T,i},i}$; order these to obtain $\hat{U}_{1,i}, \ldots, \hat{U}_{N_{T,i},i}$ and set $\tau_{j,i} = \hat{U}_{j,i}$, $j = 1, \ldots, N_{T,i}$, $\tau_0 = 0$ and $\tau_{N_{T,i}+1,i} = T$.
   c. simulate $Y_{1,i}, \ldots, Y_{N_{T,i},i}$ by generating $N_{T,i}$ independent random variates from the distribution $f$
   d. simulate $W_{\tau_{1,i},i}, \ldots, W_{\tau_{N_{T,i},i},i}, W_{N_{T,i}+1,i}$ by setting
      
      $W_{\tau_{j,i},i} = W_{\tau_{j-1,i},i} + N(0, \tau_j - \tau_{j-1,i})$
      
      $j = 1, \ldots, N_{T,i}, N_{T,i} + 1, W_{\tau_0,i} = 0$
   e. for $j = 1 : m_2$ (Inner loop)
      i. for $k = 1 : (N_{T,i} + 1)$
         
         A. simulate $M_{k,i,j}$ using equations 11 and 12 with $x = \gamma \tau_{k-1,i} + W_{\tau_{k-1,i},i} + \sum_{l=1}^{N_{T,i}} Y_{l,i}, l = \tau_{k,i} - \tau_{k-1,i}$ and $\tau_0 = 0$
         ii. end
         iii. compute $Q_{i,j} = \exp(-rT) \max (B, M_{1,i,j}, \ldots, M_{N_{T,i},i,j}, M_{N_{T,i}+1,i,j})$
      f. end
   g. $P_i = \frac{1}{m_2} \sum_{j=1}^{m_2} Q_{i,j}
2. end
3. $I = \frac{1}{m_1} \sum_{i=1}^{m_1} P_i$
4. $V = \frac{1}{m_1} \left( \sum_{i=1}^{m_1} \left[ \sum_{j=1}^{m_2} Q_{i,j} - \sum_{j=1}^{m_2} Q_{i,j}^2 \right] - \left[ \sum_{i=1}^{m_1} \sum_{k=1}^{m_2} P_i P_k - \sum_{i=1}^{m_1} P_i^2 \right] \cdot \frac{1}{m_1 (m_1 - 1)} \sum_{i=1}^{m_2} (Q_{i,j} - P_i)^2 \right) + \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2 (m_2 - 1)} \sum_{j=1}^{m_2} (Q_{i,j} - P_i)^2$

We use this algorithm in chapter 5 to price a lookback put option in the Kou model and to compute variances of the estimator of the price. In section 5.1, the assumptions underlying the Kou model are recalled and numerical results are presented in section 5.2.

5 Numerical Experiments

In this section, we firstly review the Kou model, [7] and [8], and parameterize the model. Secondly, we price a lookback put option in the Kou model using Algorithm 4.1, investigate the order of magnitude of the variance of estimator (13) and compare the results to those produced by algorithm 3.1.
5.1 The Kou Model

The Kou Model was introduced in [7] and [8]. Following [8], the relevant jump-diffusion process is defined as follows.

Definition 5.1 Jump-diffusion Process for the Kou Model:

$(N_t)$ is a Poisson Process, with intensity $\lambda$ and jump times $\tau$, $(W_t)$ is a Brownian Motion and $\{Y_i\}$ are independent, identically distributed random variables with distribution $f$.

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$$  \hspace{1cm} (15)

where

$$f(dx) = \begin{cases} p \eta_1 \exp(-\eta_1 x) 1_{x>0} & \text{if } \frac{\eta_1}{\eta_1 - 1} + (1-p) \frac{\eta_2}{\eta_2 + 1} < 1 \\ (1-p) \eta_2 \exp(\eta_2 x) 1_{x<0} & \text{if } \frac{\eta_1}{\eta_1 - 1} + (1-p) \frac{\eta_2}{\eta_2 + 1} > 1 \\ \end{cases}$$

$$\gamma = r - \frac{1}{2} \sigma^2 - \lambda \left( p \frac{\eta_1}{\eta_1 - 1} + (1-p) \frac{\eta_2}{\eta_2 + 1} - 1 \right)$$

and $0 \leq p \leq 1, \lambda > 0, \eta_1 > 1, \eta_2 > 0, \sigma > 0$.

As in section 2.1, we use equation (15) to define the stock price process.

Definition 5.2 Stock price Process for the Kou model:

Let $X_t$ be defined as in Definition 5.1. Then the stock price process is given by

$$S_t = S_0 \exp(X_t)$$  \hspace{1cm} (16)

In the next section, we will parameterize this model and price lookback options.

5.2 Pricing Lookback options

Given the share price process (16), we now parameterize the Kou model as in [8], by setting $r = 0.05$, $\sigma = 0.2, p = 0.3, \frac{1}{\eta_1} = 0.02, \frac{1}{\eta_2} = 0.04, \lambda = 3$ and $S_0 = 100$. The lookback put option furthermore uses the parameters $B = 110$ and $T = 1$. The true price of the option, as taken from [8], is 17.01.

We now produce a table which presents prices and variances and observe that prices converge to the correct value and variances have the suggested orders of magnitude. Table 1 uses algorithm 4.1 to estimate prices of lookback put options and variances of the estimators, for different values of $m_1$ and $m_2$.

Table 1 verifies Lemma 3.9 as increasing $m_1$ by a factor of 2 reduces the variance by a factor of 2, whereas increasing $m_2$ by a factor of 2 does not have this effect.

<table>
<thead>
<tr>
<th>Price</th>
<th>Variance</th>
<th>$m_1 = 2000$</th>
<th>$m_1 = 4000$</th>
<th>$m_1 = 8000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_2 = 2000$</td>
<td>16.54651</td>
<td>16.8703</td>
<td>17.0211</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1015</td>
<td>0.0517</td>
<td>0.0269</td>
<td></td>
</tr>
<tr>
<td>$m_2 = 4000$</td>
<td>17.5131</td>
<td>17.2341</td>
<td>17.0985</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1088</td>
<td>0.0567</td>
<td>0.0273</td>
<td></td>
</tr>
<tr>
<td>$m_2 = 8000$</td>
<td>16.8274</td>
<td>16.6758</td>
<td>16.9083</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1039</td>
<td>0.0497</td>
<td>0.0263</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Estimates of Prices and of Variances of estimators of lookback put options based on Algorithm 4.1

The Table 2 produces estimates of prices and estimates of variances of estimators of prices of lookback put options based on Algorithm 3.1.

We see that the variances are order of magnitude $O(m_1)$, as indicated by Lemma 3.1 and increasing $m_1$ and $m_2$ leads to improvements in the estimates of prices. We remark that parameters
Table 2: Estimates of Prices and of Variances of estimators of lookback put options based on Algorithm 3.1

<table>
<thead>
<tr>
<th></th>
<th>(m_1 = 2000)</th>
<th>(m_1 = 4000)</th>
<th>(m_1 = 8000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m_2 = 2000)</td>
<td>17.0637</td>
<td>16.8584</td>
<td>16.8787</td>
</tr>
<tr>
<td></td>
<td>0.1158</td>
<td>0.0575</td>
<td>0.0290</td>
</tr>
<tr>
<td>(m_2 = 4000)</td>
<td>16.9203</td>
<td>17.1792</td>
<td>16.8999</td>
</tr>
<tr>
<td></td>
<td>0.1212</td>
<td>0.0598</td>
<td>0.0290</td>
</tr>
<tr>
<td>(m_2 = 8000)</td>
<td>16.9061</td>
<td>16.9339</td>
<td>16.9466</td>
</tr>
<tr>
<td></td>
<td>0.1223</td>
<td>0.0599</td>
<td>0.0287</td>
</tr>
</tbody>
</table>

For Algorithm 4.1, \(m_2\) determines the number of points used to approximate the inner expectation in equation (7) and \(m_1\) determines the number of inner expectations over which we average, so \(m_1\) and \(m_2\) determine the total number of points used in the approximation. Regarding Algorithm 3.1, the total number of points used is \(m_1\), each of which is of dimension \(m_2\). It seems fair to say that variances seem to be similar magnitude, even though the parameters \(m_1\) and \(m_2\) for the respective algorithms do not correspond exactly.

However, it is recalled that for fixed \(m_2\), Algorithm 3.1 produces biased estimates and that one converges to the biased estimate at a rate of \(O\left(m_1^{-1}\right)\). It is not known a priori though how large this bias is. Algorithm 4.1 produces unbiased estimates for prices of lookback options and the variance of the estimator depends on \(m_1\) and \(m_2\) and is \(O\left(m_1^{-1}\right)\) for fixed \(m_2\).

Finally, it should be pointed out that we applied our methodology to the Kou model for purposes of illustration and to investigate the order of magnitude of the variance of estimator (13). The methods for pricing lookback options in the Kou model presented in [8] are more efficient wrt CPU time than the methods presented in this paper. However, our methodology is more general, the distribution of the jump sizes does not need to be an asymmetric exponential distribution, it can be any distribution one can simulate from using Monte Carlo methods.

### 6 Conclusion

In this paper, a method to estimate prices of lookback options in jump-diffusion models without bias was introduced, analyzed and investigated numerically. The method is very general, it can be applied to any jump-diffusion model, given that one can simulate the magnitudes of the jumps using Monte Carlo methods.

The above methodology allows to simulate jump-diffusion processes. Of course it cannot be extended to the more general class of Lévy processes. Lévy processes can exhibit an infinite number of jumps and hence cannot be simulated using Algorithm 4.1, which crucially relies on jump-diffusions having a finite number of jumps. However, some infinite activity processes may be approximated using jump-diffusions, e.g. [1] and [5], in which case the above methodology would be useful again.

Further interesting areas of research would be to investigate if the above convergence rates of variances can be improved on, e.g. using variance reduction techniques, see e.g. [6]. The author will also investigate if Quasi-Monte Carlo methods, see e.g. [4], allow to achieve faster rates of convergence than the Monte Carlo Methods presented here.
A Proofs of Lemma 3.8 and Proposition 3.2

In appendix A, we will provide the proofs of Lemma 3.8 and Proposition 3.2. We firstly present the proof of Lemma 3.8.

**Lemma 3.8**  The variance of the estimator (13) is given by the following formula

\[
\text{Var}(I) = \frac{\text{Var}(f_1(N_T, \tau, Y, W))}{m_1} + \frac{1}{m_1m_2} \left( \mathbb{E}[f_2(N_T, \tau, Y, W)] - \mathbb{E}[f_1(N_T, \tau, Y, W)]^2 \right)
\]

**Proof.** We introduce the following notation \( G = \sigma(N_{T,i}, \tau_i, Y_i, W_i, i = 1, \ldots, m_1), G_i = \sigma(N_{T,i}, \tau_i, Y_i, W_i, i = 1, \ldots, m_1) \). Clearly,

\[
\text{Var}[I] = \text{Var}[\mathbb{E}[I | G]] + \mathbb{E}[\text{Var}[I | G]]
\]

We firstly compute \( \text{Var}[\mathbb{E}[I | G]] \)

\[
\mathbb{E} \left[ \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) | G \right] = \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} \mathbb{E}[g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) | G]
\]

\[
= \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} \mathbb{E}[g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) | G_i] = \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} f_1(N_{T,i}, \tau_i, Y_i, W_i) = \frac{1}{m_1} \sum_{i=1}^{m_1} f_1(N_{T,i}, \tau_i, Y_i, W_i)
\]

Consequently

\[
\text{Var}[\mathbb{E}[I | G]] = \frac{1}{m_1} \sum_{i=1}^{m_1} \text{Var}(f_1(N_{T,i}, \tau_i, Y_i, W_i)) = \frac{\text{Var}(f_1(N_T, \tau, Y, W))}{m_1}
\]

assuming \( \text{Var}(f_1(N_T, \tau, Y, W)) \) is well-defined. Equation (18) follows as \( N_{T,i}, Y_i, \tau_i, W_i \) and \( N_{T,j}, Y_j, \tau_j, W_j \) are independent for \( i \neq j \), equation (19) follows Lemma 3.7, equation (20) follows as \( N_{T,i}, Y_i, \tau_i, W_i, i = 1, \ldots, m_1 \) are independently distributed and equation (21) follows because \( N_{T,i}, Y_i, \tau_i, W_i, i = 1, \ldots, m_1 \) are identically distributed. We now compute \( \mathbb{E}[\text{Var}[I | G]] \)

\[
\text{Var} \left[ \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) | G \right] = \frac{1}{(m_1m_2)^2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \text{Var}(g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) | G)
\]

The above follows as conditional on \( G \), for fixed \( i \), \( g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}), j = 1, \ldots, m_2 \) are independent and for fixed \( i, k, i \neq k \), \( g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) \) and \( g(N_{T,k}, \tau_k, Y_k, W_k, M_{k,j}), j = 1, \ldots, m_2 \)
and \( l = 1, \ldots, m_2 \) are independent.

\[
\frac{1}{(m_1 m_2)^2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \text{Var} \left( g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) \mid \mathcal{G} \right)
\]

\[
= \frac{1}{(m_1 m_2)^2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \left( \mathbb{E} \left[ g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j})^2 \mid \mathcal{G} \right] - \left( \mathbb{E} [g(N_{T,i}, \tau_i, Y_i, W_i, M_{i,j}) \mid \mathcal{G}] \right)^2 \right)
\]

\[
= \frac{1}{(m_1 m_2)^2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \left( f_2(N_{T,i}, \tau_i, Y_i, W_i) - (f_1(N_{T,i}, \tau_i, Y_i, W_i))^2 \right)
\]

Equation (22) follows as \( N_{T,i}, Y_i, \tau_i, W_i \) and \( N_{T,j}, Y_j, \tau_j, W_j \) are independent for \( i \neq j \) and equation (23) follows from the Lemma 3.7. Hence

\[
\mathbb{E} \left[ \text{Var} \left( I \mid \mathcal{G} \right) \right] = \frac{1}{(m_1)^2 m_2} m_1 \left[ \mathbb{E} [f_2(N_{T}, \tau, Y, W)] - \mathbb{E} \left[ (f_1(N_{T}, \tau, Y, W))^2 \right] \right]
\]

Equation (24) follows as \( N_{T,i}, Y_i, \tau_i, W_i, i = 1, \ldots, m_1 \) are identically distributed.

We finally present the proof of Proposition 3.2.

**Proposition 3.2** Let

\[ J = V_1 + V_2 \]

where

\[ V_1 = \frac{1}{m_1} \left( \frac{\sum_{i=1}^{m_1} \left[ \sum_{j=1}^{m_2} g_{ij} g_k - \sum_{j=1}^{m_2} \bar{g}_k \right]}{m_2 (m_2 - 1) m_1} - \left( \frac{\sum_{i=1}^{m_1} \sum_{k=1}^{m_1} \bar{g}_i \bar{g}_k - \sum_{i=1}^{m_1} \bar{g}_i^2}{m_1 (m_1 - 1)} \right) \right) \]

and

\[ V_2 = \frac{1}{m_2} \sum_{i=1}^{m_1} \frac{1}{m_2 (m_2 - 1)} \sum_{j=1}^{m_2} \left( g_{ij} - \bar{g}_i \right)^2 \]

and \( g_{ij} = g(N_{T,i}, \tau_i, Y_i, W_i, M_{ij}) \), \( i = 1, \ldots, m_1, j = 1, \ldots, m_2 \), \( \bar{g}_i = \sum_{j=1}^{m_2} \frac{g_{ij}}{m_2}, i = 1, \ldots, m_1 \) and \( \bar{g} = \sum_{i=1}^{m_1} \frac{\bar{g}_i}{m_1} \). Then \( J \) is an unbiased estimator of \( \text{Var} (I) \).

**Proof.** Functions \( f_1 \) and \( f_2 \) are defined as in Lemma 3.8.

We firstly compute \( \mathbb{E} [V_1] \)

\[
V_1 = \frac{1}{m_1} \left( \frac{\sum_{i=1}^{m_1} \left[ \sum_{j=1}^{m_2} g_{ij} g_k - \sum_{j=1}^{m_2} \bar{g}_k \right]}{m_2 (m_2 - 1) m_1} - \left( \frac{\sum_{i=1}^{m_1} \sum_{k=1}^{m_1} \bar{g}_i \bar{g}_k - \sum_{i=1}^{m_1} \bar{g}_i^2}{m_1 (m_1 - 1)} \right) \right)
\]

\[
= \frac{1}{m_1} \left( \frac{\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} g_{ij} g_k}{m_2 (m_2 - 1) m_1} - \frac{1}{m_1} \sum_{i=1}^{m_1} \frac{\bar{g}_i \bar{g}_k}{m_1 (m_1 - 1)} \right)
\]
Firstly,

\[
E \left[ \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} g_{ij} \bar{g}_k \right] = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} E \left[ g_{ij} \bar{g}_k \right]
\]

\[
= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \sum_{k \neq j} E \left[ g_{ij} \bar{g}_k \mid G_i \right]
\]

\[
= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \sum_{k \neq j} E \left[ g_{ij} \mid G_i \right] E \left[ \bar{g}_k \mid G_i \right]
\]

\[
= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \sum_{k \neq j} E \left[ f_1 \left( N_{T, i}, \tau_i, Y_i, W_i \right) \right]
\]

\[
= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \sum_{k \neq j} E \left[ f_1 \left( N_{T, i}, \tau_i, Y_i, W_i \right) \right]
\]

\[
= m_1 m_2 \left( m_2 - 1 \right) E \left[ f_1^2 \left( N_T, \tau, Y, W \right) \right]
\]

Secondly,

\[
E \left[ \sum_{i=1}^{m_1} \sum_{k \neq i} g_{i} \bar{g}_k \right] = \sum_{i=1}^{m_1} \sum_{k \neq i} E \left[ g_{i} \bar{g}_k \right]
\]

\[
= \sum_{i=1}^{m_1} \sum_{k \neq i} E \left[ g_{i} \mid G_i \right] E \left[ \bar{g}_k \mid G_i \right]
\]

\[
= \sum_{i=1}^{m_1} \sum_{k \neq i} \left( \sum_{j=1}^{m_2} E \left[ g_{ij} \mid G_i \right] \right) \left( \sum_{j=1}^{m_2} E \left[ g_{kj} \mid G_i \right] \right)
\]

\[
= \sum_{i=1}^{m_1} \sum_{k \neq i} \left( \sum_{j=1}^{m_2} E \left[ f_1 \left( N_{T, i}, \tau_i, Y_i, W_i \right) \right] \right) \left( \sum_{j=1}^{m_2} E \left[ f_1 \left( N_{T, i}, \tau_i, Y_i, W_i \right) \right] \right)
\]

\[
= \sum_{i=1}^{m_1} \sum_{k \neq i} \left( E \left[ f_1 \left( N_{T, i}, \tau_i, Y_i, W_i \right) \right] \right) \left( E \left[ f_1 \left( N_{T, i}, \tau_i, Y_i, W_i \right) \right] \right)
\]

\[
= m_1 \left( m_1 - 1 \right) E \left[ f_1 \left( N_T, \tau, Y, W \right) \right] E \left[ f_1 \left( N_T, \tau, Y, W \right) \right]
\]

\[
= m_1 \left( m_1 - 1 \right) \left( E \left[ f_1 \left( N_T, \tau, Y, W \right) \right] \right)^2
\]

hence

\[
E \left[ V_1 \right] = \frac{\text{Var} \left[ f_1 \left( N_T, \tau, Y, W \right) \right]}{m_1}
\]

which completes the first part of the proof.

We now compute $E \left[ V_2 \right]$. Firstly,

\[
(g_{ij} - \bar{g}_i)^2 = g_{ij}^2 - 2g_{ij} \bar{g}_i + (\bar{g}_i)^2
\]

\[
= g_{ij}^2 - 2g_{ij} \frac{\sum_{k=1}^{m_2} g_{ik}}{m_2} + \left( \frac{\sum_{k=1}^{m_2} g_{ik}}{m_2} \right)^2
\]
Hence

\[
\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (g_{ij} - \bar{g}_i)^2
\]

\[
= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} g_{ij}^2 - 2 \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} g_{ij} \sum_{k=1}^{m_2} \frac{g_{ik}}{m_2} + \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \left( \frac{g_{ik}}{m_2} \right)^2
\]

\[
= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} g_{ij}^2 - 2 \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} g_{ij} \left( \sum_{k=1}^{m_2} \frac{g_{ik}}{m_2} \right) + \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \left( \sum_{k=1}^{m_2} \frac{g_{ik}}{m_2} \right)
\]

Hence

\[
\mathbb{E} \left[ \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (g_{ij} - \bar{g}_i)^2 \right] = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \mathbb{E} [g_{ij}^2] - \sum_{i=1}^{m_1} \mathbb{E} \left[ \sum_{j=1}^{m_2} \left( \sum_{k=1}^{m_2} \frac{g_{ik}}{m_2} \right) \right]
\]

Firstly,

\[
\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \mathbb{E} [g_{ij}^2] = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \mathbb{E} [g_{ij}^2 | G_i]
\]

\[
= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \mathbb{E} [f_2 (N_{T,i}, \tau_i, Y_i, W_i)] = m_2 \sum_{i=1}^{m_1} \mathbb{E} [f_2 (N_{T,i}, \tau_i, Y_i, W_i)] = m_2 m_1 \mathbb{E} [f_2 (N_{T,i}, \tau_i, Y_i, W_i)]
\]

Secondly,

\[
\sum_{i=1}^{m_1} \mathbb{E} \left[ \sum_{j=1}^{m_2} g_{ij} \sum_{k=1}^{m_2} g_{ik} \right] = \sum_{i=1}^{m_1} \frac{1}{m_2} \mathbb{E} \left[ \sum_{j=1}^{m_2} g_{ij}^2 + \sum_{j=1}^{m_2} \sum_{k \neq j} g_{ij} g_{ik} \right]
\]

\[
= \frac{1}{m_2} \sum_{i=1}^{m_1} \left( \sum_{j=1}^{m_2} \mathbb{E} [f_2 (N_{T,i}, \tau_i, Y_i, W_i)] + \sum_{j=1}^{m_2} \sum_{k \neq j} \mathbb{E} [g_{ij} g_{ik} | G_i] \mathbb{E} [g_{ik} | G_i] \right)
\]

\[
= \frac{1}{m_2} \sum_{i=1}^{m_1} \left( m_2 \mathbb{E} [f_2 (N_{T,i}, \tau_i, Y_i, W_i)] + \sum_{j=1}^{m_2} \sum_{k \neq j} \mathbb{E} [f_1 (N_{T,i}, \tau_i, Y_i, W_i) f_1 (N_{T,i}, \tau_i, Y_i, W_i)] \right)
\]

\[
= m_1 \mathbb{E} [f_2 (N_T, \tau, Y, W)] + m_1 (m_2 - 1) \mathbb{E} [f_2^2 (N_T, \tau, Y, W)]
\]

It now follows that

\[
\mathbb{E} \left[ \sum_{i=1}^{m_1} \frac{1}{m_2} \sum_{j=1}^{m_2} (g_{ij} - \bar{g}_i)^2 \right]
\]

\[
= \frac{1}{m_2} \sum_{i=1}^{m_1} \frac{1}{m_2} (m_2 - 1) \sum_{j=1}^{m_2} (g_{ij} - \bar{g}_i)^2
\]

\[
= \frac{1}{m_2} \sum_{i=1}^{m_1} \frac{1}{m_2} (m_2 - 1) \mathbb{E} [f_2 (N_T, \tau, Y, W)] - (m_1 \mathbb{E} [f_2 (N_T, \tau, Y, W)] + m_1 (m_2 - 1) \mathbb{E} [f_2^2 (N_T, \tau, Y, W)])
\]

\[
= \frac{1}{m_2} \sum_{i=1}^{m_1} \frac{1}{m_2} (m_2 - 1) \mathbb{E} [f_2 (N_T, \tau, Y, W)] - m_1 (m_2 - 1) \mathbb{E} [f_2^2 (N_T, \tau, Y, W)]
\]

\[
= \frac{1}{m_1 m_2} \left[ \mathbb{E} [f_2 (N_T, \tau, Y, W)] - \mathbb{E} [f_2^2 (N_T, \tau, Y, W)] \right]
\]
Hence
\[ \mathbb{E}[J] = \text{Var}(I) \]
using Lemma 3.8, which completes the proof.

References


